

## REMARKS

Claims 6-9, 26, 27 and 32-34 are pending in this application. Claim 34 is amended herein to clarify and more particularly define the invention. No new matter is added by this amendment. In light of the following amendments and remarks, applicants respectfully request reconsideration of this application and allowance of the claims to issue.

### **I. Rejection under 35 U.S.C. § 112**

A. The Action states that claims 6-9, 26, 27 and 32-34 stand rejected under 35 U.S.C. § 112, first paragraph, for allegedly failing to comply with the written description requirement.

Claim 6 as presented herein encompasses a specific genus of nucleic acid sequences encoding peptides immunochemically reactive with antibodies to the Epstein Barr Virus (EBV) VCA-p18 or VCA-p40 proteins, comprising an epitope of the VCA-p18 or VCA-p40 protein, encoded within the EBV open reading frames BFRF3 and BdRF1, respectively and wherein said antibodies are antibodies having the same reactivity with VCA-p18 as antibodies produced by the hybridomas deposited at the European Collection of Animal Cell Cultures under deposit numbers 93020413 or 93020412 or antibodies having the same reactivity with VCA-p40 as antibodies produced by the hybridoma deposited at the European Collection of Animal Cell Cultures under deposit number 93020414.

In claim 6, the nucleic acid sequences encode peptides comprising an epitope of the VCA-p18 or VCA-p40 protein. The peptides encoded by the nucleic acid sequences of claim 6 are further defined by their immunoreactivity with hybridoma-derived antibodies to VCA-p18 and VCA-p40 defined as European Collection of Animal Cell Cultures deposit numbers 93020413 or 93020412 (VCA-p18) or deposit number 93020414 (VCA-p40). As one of skill in the art would recognize, hybridoma derived antibodies are monoclonal. One of skill in the art further recognizes that monoclonal antibodies are specific to a single epitope and that all of the antibodies produced from a single hybridoma are identical. Therefore, contrary to the arguments in the Office Action,

these antibodies are not heterogeneous and, thus, are sufficient to define the genus of nucleic acid sequences encompassed by claim 6.

Furthermore, the specification demonstrates actual reduction to practice of the nucleic acids of claim 6 and in particular, provides several examples of peptides comprising epitopes of this invention (e.g., SEQ ID NOs:2, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and 22; see pages 7-13 for description of peptides and fragments of this invention and Examples 4 and 5, Figures 4-6 and Table 1) that are reactive with the EBV VCA-p18 or VCA-p40 monoclonal antibodies of claim 6 and also provides examples of nucleic acid sequences encoding such peptides (e.g., SEQ ID NO:1, SEQ ID NO:3). Accordingly, one of skill in the art would recognize that applicants were in possession of the nucleic acid sequences of claim 6 at the time the present application was filed, as evidenced by the large numbers of representative species disclosed in the specification.

Therefore, applicants respectfully submit that all of the members of the genus of nucleic acids of claim 6 are adequately defined both structurally and functionally, leading one of ordinary skill in the art to the reasonable conclusion that applicants were in possession of the invention of claim 6 at the time this application was filed.

Claims 7 and 8 as presented herein respectively encompass a specific genus of nucleic acid sequences comprising the nucleotide sequence or a subsequence of SEQ ID NO:1, wherein the subsequence encodes a peptide that comprises an epitope that is immunochemically reactive with antibodies to EBV VCA-p19 protein (claim 7), and a specific genus of nucleic acid sequences comprising the nucleotide sequence or a subsequence of SEQ ID NO:3, wherein the subsequence encodes a peptide that comprises an epitope that is immunochemically reactive with antibodies to EBV VCA-p40 protein (claim 8).

In both claims 7 and 8, the subsequence is defined as encoding a peptide comprising an epitope. It would be readily recognized by one of skill in the art that applicants were in

possession of the genus of nucleic acid sequences of claims 7 and 8 at the time the application was filed because the specification provides several examples of peptides comprising epitopes of this invention (e.g., SEQ ID NOs: 2, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 and 22; see pages 7-13 for description of peptides and fragments of this invention and Examples 4 and 5, Figures 4-6 and Table 1) that are reactive with the EBV VCA-p18 or VCA-p40 antibodies of this invention and also provides examples of nucleic acid sequences encoding such peptides (e.g., SEQ ID NO:1, SEQ ID NO:3). Thus, one of skill in the art would recognize that applicants were in possession of the nucleotide sequences and subsequences of claims 7 and 8 at the time the present application was filed, as evidenced by the large numbers of representative species disclosed in the specification.

In particular, all of the members of the genus of the nucleotide sequences of claims 7 and 8 could be readily identified by one of ordinary skill in the art on the basis of the disclosure of the nucleotide sequences of SEQ ID NO:1 or SEQ ID NO:3. Such a genus is not overly broad, considering that every member must be a subsequence of a disclosed sequence (SEQ ID NO:1 or SEQ ID NO:3), thereby defining the members of the genus structurally AND every member of the genus must also meet the functional requirement of encoding an EBV peptide comprising an epitope that is immunochemically reactive with antibodies to the EBV VCA-p18 protein or the EBV VCA-p40 protein. Thus, all of the members of the genus of nucleic acids of claims 7 and 8 are adequately defined both structurally and functionally, leading one of ordinary skill in the art to the reasonable conclusion that applicants were in possession of the invention of claims 7 and 8 at the time this application was filed.

Claims 9, 26 and 27 depend from claims 6, 7 and 8, respectively, and recite a vector molecule comprising the nucleic acid molecule of each respective independent claim. Because the nucleic acid sequences of claims 6, 7 and 8 are adequately described in the specification, the vectors of these claims are adequately described as well.

With regard to claims 32-33, the specification presents data that demonstrate that the inventors produced more than 330 12 mers of VCA-p40 and more than 160 12 mers of VCA-p18 as described in Examples 4 and 5 and as shown in Figures 4-6 and in Table 1 of the specification, thereby adequately describing the genus of 12 contiguous amino acids as set forth in these claims.. Specifically, Example 4 describes the production of a full set of peptides with a length of 12 amino acids and an overlap of 11 amino acids of the amino acid sequences of both ORFs BFRF3 (VCA-p18) and BDRF1 (VCA-p40) (page 30). These peptides were assayed for immunoreactivity with EBV-specific antibodies (Example 4, page 31 and Example 5, page 33) and results of these assays are shown for the VCA-p18 peptides in Figures 4 and 5 and for the VCA-p40 peptides in Figure 6. Specifically, Figure 6 shows immunoreactivity results of almost 340 peptides of VCA-p40 and Figures 4 and 5 show such results for more than 160 peptides of VCA-p18. Thus, one of skill in the art would reasonably conclude that the peptides of claims 32 and 33 are adequately supported in the specification

Further, claim 34 as presented herein recites an isolated nucleic acid sequence encoding the amino acid sequence of SEQ ID NO:5 or SEQ ID NO:6 or a combination of both, wherein said amino acid sequence is immunochemically reactive with antibodies to the Epstein-Barr Virus VCA-p18 protein. Thus, claim 34 recites a nucleic acid sequence encoding the specific amino acid sequences of SEQ ID NO: 5 and/or SEQ ID NO:6, which are disclosed in the specification at least on page 9, second paragraph. Thus, the nucleic acid sequences of claim 34 are adequately supported in the specification.

Thus, at least for the reasons set forth above, applicants believe that this rejection has been overcome and its withdrawal and allowance of the pending claims are respectfully requested.

**B.** The Action states that claims 6-9, 26, 27 and 32-34 stand rejected under 35 U.S.C. § 112, first paragraph, for allegedly failing to comply with the enablement requirement. Specifically, the Action states that to the extent that the claimed sequences are not adequately

described in the instant disclosure, claims 6-9, 26, 27 and 32-34 are also rejected under 35 U.S.C. § 112, first paragraph, as allegedly containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to make and/or use the invention.

As discussed above, the subject matter of claims 6-9, 26, 27 and 32-34 is adequately described in the present specification. The specification not only adequately discloses the full genus of nucleic acid sequences of this invention, but also provides detailed teachings of how to make and use these nucleic acid sequences. See, in particular, the Examples set forth on pages 22-26, wherein numerous working examples are provided of the production and testing of numerous peptides of this invention. Thus, applicants respectfully submit that the present invention is adequately enabled and applicants thereby respectfully request withdrawal of this rejection.

## **II. Rejection under 35 U.S.C. § 102(b)**

A. The Action states that claims 6-9, 26, 27 and 32-34 stand rejected under 35 U.S.C. § 102(b) as allegedly anticipated by Laux et al. (*EMBO J.* 7:769-774 (1988)). Specifically, the Action states that Laux et al. teaches a nucleic acid sequence comprising instant SEQ ID NO:1 which encodes at least 12 contiguous amino acids of EBV VCA-p18 (the amino acid sequence SEQ ID NO:5). The Action further states that Laux et al. teaches a nucleic acid sequence comprising a sequence that shares 98.8% homology with instant SEQ ID NO:3 (subsequence thereof), which encodes 12 contiguous amino acids of an EBV VCA-40. On this basis, the Action concludes that Laux et al. anticipates the instant claims. Applicants respectfully disagree and traverse this rejection.

Specifically, applicants have performed multiple alignments comparing both the nucleotide and amino acid sequences of the present invention with those of Figure 2 of Laux et al. (NCBI Accession No. Y00835.1) and the sequence homology asserted in the Action to be present was not found (See enclosed Alignments 1-6). Accordingly, applicants respectfully submit that Laux et al. fails to disclose the nucleotide sequences of SEQ ID NO:1, or SEQ ID

NO:3 or any subsequences thereof encoding 12 contiguous amino acids, as claimed herein. If the Examiner maintains this rejection, it is respectfully requested that the Examiner specifically point out what portion of the sequences of Laux et al. have homology with the sequences of the present invention. Otherwise, applicants respectfully request that this rejection be withdrawn.

**B.** The Action states that claims 6-9, 26, 27 and 32-34 stand rejected under 35 U.S.C. § 102(b) as allegedly anticipated by Bankier et al. (*Mol. Biol. Med.* 1:425-445 (1983)). Specifically, the Action states that Bankier et al. teaches a nucleic acid sequence comprising instant SEQ ID NO:1, which encodes at least 12 contiguous amino acids of EBV VCA-p18 (the amino acid sequence SEQ ID NO:5). The Action further states that Bankier et al. teaches a nucleic acid sequence comprising a sequence that shares 98.8% homology with instant SEQ ID NO:3 (subsequence thereof), which encodes 12 contiguous amino acids of an EBV VCA-40. On this basis, the Action concludes that Bankier et al. anticipates the instant claims. Applicants respectfully disagree and traverse this rejection.

Specifically, applicants have performed multiple alignments comparing both the nucleotide and amino acid sequences of the present invention with those of Figure 2 of Bankier et al. and the sequence homology asserted in the Action to be present was not found (See enclosed Alignments 7-38). Accordingly, applicants respectfully submit that Bankier et al. fails to disclose the nucleotide sequences of SEQ ID NO:1, or SEQ ID NO:3 or any subsequences thereof encoding 12 contiguous amino acids as claimed herein. If the Examiner maintains this rejection, it is respectfully requested that the Examiner specifically point out what portion of the sequences of Bankier et al. have homology with the sequences of the present invention. Otherwise, applicants respectfully request that this rejection be withdrawn.

The points and concerns raised in the outstanding Office Action having been addressed in full, it is respectfully submitted that all of the claims of this application are in condition for allowance, which action is respectfully requested. Should the Examiner have any remaining

concerns, the Examiner is invited and encouraged to contact the undersigned attorney directly by telephone in order to expedite the prosecution of this application to allowance.

The Commissioner is authorized to charge Deposit Account No. 50-0220 in the amount of \$120.00 as fee for a one-month extension of time. This amount is believed to be correct. However, the Commissioner is authorized to charge any deficiency or credit any overpayment to Deposit Account No. 50-0220.

Respectfully submitted,



Mary L. Miller  
Registration No. 39,303

**Customer No. 20792**

Myers Bigel Sibley & Sajovec, P.A.  
P. O. Box 37428  
Raleigh, North Carolina 27627  
Telephone: (919) 854-1400  
Facsimile: (919) 854-1401

CERTIFICATE OF EXPRESS MAILING UNDER 37 CFR 1.10

**"Express Mail" mailing label number: EV887524629US**

Date of Deposit: July 23, 2007

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

  
Tracy Wallace

9310-13DVCTDV SEQ ID NO 1.xdna x Laux et al. EBV terminal gene.xdna => DNA Parallel

DNA sequence 538 bp catgatggcacg ... aaacagtagccc linear

DNA sequence 2227 bp gcagtgtgtgaa ... aaaaaaaaaa linear

Method: Blocks (Martinez)  
Layout: Standard  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Translation: Off

Alignment 1. Comparison of nucleotide sequence  
of SEQ ID NO:1 with the nucleotide sequence of  
Fig. 2 of Laux et al.

```

-----
1 gcagtgtgtgaagattgtcacagctgctggtttggagaaaacgggggtgggcggtgatca 60
   .               20               .               40               .               60
-----
61 gggagaacaattccccggggacacctgcacgagaccctgggctctcaggaactccgccc 120
   .               80               .               100              .               120
-----
121 aggtcttgccaattggggtgatcctgtagcgccgcgggtttcagcatcacagggtatttttg 180
   .               140              .               160              .               180
-----
181 cctgaagcttgctggggcgtaaataccctctcgccctgtttctcagagagcatttcaggcc 240
   .               200              .               220              .               240
-----
241 ggttttgcagtcgctgctgcagctatggggtcctagaaatggtgccaatgggcgcgggt 300
   .               260              .               280              .               300
-----
301 cccctagccccggcggggatccggatgggtacgatggcggaacaactcccaatatcca 360
   .               320              .               340              .               360
-----
361 tctgcttctggctcttcttggaacacccccacccacccaacgatgaggaacgtgaatct 420
   .               380              .               400              .               420
-----
421 aatgaagagccccacccgcttatgaggaccatattggggcaatggcgaccgtcactcg 480
   .               440              .               460              .               480
-----
481 gactatcaaccactaggaacccaagatcaaagtctgtacttgggattgcaacacgacggg 540
   .               500              .               520              .               540
-----
541 aatgacgggctccctcccccctccctactctccacgggatgactcatctcaacacatatac 600
   .               560              .               580              .               600

```



-----  
 601 gaagaagcgggcagaggaagtatgaatccagtatgcctgcctgtaattgttgccgacctac 660  
                                 620                                640                                660  
 -----

661 ctcttttggctggcggctattgcccgcctcgtgtttcacggcctcagttagtagtaccgttgtg 720  
                                 680                                700                                720  
 -----

721 accgccaccggccttggccctctcacttctactcttggcagcagtgggccagctcatatgcc 780  
                                 740                                760                                780  
 -----

781 gctgcacaaaggaaactgctgacaccgggtgacagtgcttactgcgggttgctactttcttt 840  
                                 800                                820                                840  
 -----

841 gcaatttgcctaacatggaggattgaggacccaccttttaattctcttctgtttgcattg 900  
                                 860                                880                                900  
 -----

901 ctggccgcagctggcggactacaaggcatttacgttcttggtgatgcttgtgctcctgata 960  
                                 920                                940                                960  
 -----

961 cttagcgtacagaaggagatggcgccggttgactggttggtggcgcatcatgtttttggca 1020  
                                 980                                1000                                1020  
 -----

1021 tgtgtacttgtcctcatcgtcgacgctgttttgacagctgagtcacctccttggagctgta 1080  
                                 1040                                1060                                1080  
 -----

1                   catg----- 4  
                   ||||  
 1081 actgtggtttccatgacgctgctgctactggctttcgctcctctggtctctcttcgccaggg 1140  
                                 1100                                1120                                1140  
 -----

1141 ggcctaggtactcttgggtgcagcccttttaacattggcagcagctctggcactgctagcg 1200  
                                 1160                                1180                                1200  
 -----

1201 tcaactgattttgggcacacttaacttgactacaatgttccttctcatgctcctatggaca 1260  
                                 1220                                1240                                1260  
 -----

1261 cttgtggttctcctgatttgcctcttcgtgctcttcatgtccactgagcaagatccttctg 1320  
                                 1280                                1300                                1320  
 -----

1321 gcaagactgttcctatatgctctcgcactcttgttgctagcctccgcgctaatacgtgggt 1380  
                                 1340                                1360                                1380  
 -----

-----

.....

[illegible]

.....

.....

-----atggc

Downloaded from <http://ajphaphysocpharm.sagepub.com> at National Archive Publishing Co on June 11, 2015

cg=====

80	.	100	.	120
----	---	-----	---	-----

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

=====gatgcctggcgctttgctatgaattatccaagaaacccacggagcagg

cg----tgcagag-----

1860 . 1880 . 1900 .

-----gacttttgg-----

1920                      •                      1940                      •                      1960                      •

gtgcctcggcgccaacgcgccatagacaagagggcagagagccagtgtggctggggctgg

qtgc=====

1973 | | | | | 1977

```

      280              300              320
263 gctcatgcacaccttggcgggcatccgccaccccggtccagcaggctcaggccgccgca 322
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1978 ccccttggaaca=ctt=atgtttcaagcagctcacctat=ggta==ctcaggc===== 2025
    1980              2000              2020
      340              360              380
223 tccgctgggaccggggccttggcatcatcagcgccggtccacggccgtagcccagtc-cgc 381
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2026 =====acgggtcg=cccctccgagtgaccagtcacct 2055
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      400              420
382 gacccctctgtttcttcatctattagc-----agcctccggggccgca 425
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2056 tccagactatgcatacactgaatttagcctgatattgtccccctagcc=ccggggcc==== 2110
    2060              2080              2100
      440              460              480
426 cttcgggggcgactgccgcgcctccgccgcgcagccgctcgataccgggtcaggtggcg 485
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2111 =====cagc=cctcctcagaaaactctgcatgg==agaagctg 2145
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    2120              2140
      500              520
486 ggggacaacccacgacaccgccccacgcggggcacgtaagaaacagtagccc----- 538
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
2146 gacgtgaacctc=ccccccagacctgtgtgtgtgta=tttacaacactacaataaaccca 2203
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    2160              2180              2200

```

```

-----
2204 atgtgcaaaaaaaaaaaaaaaaaa
    2220

```

2227

% Identity = 7.0 (174/2484)

///

9310-13DVCTDV SEQ ID NO 3.xdna x Laux et al. EBV terminal gene.xdna => DNA Parallel

DNA sequence 1038 bp atgctatcaggt ... cgcgtggcttga linear

DNA sequence 2227 bp gcagtgtgtgaa ... aaaaaaaaaa linear

Method: Blocks (Martinez)  
Layout: Standard  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Translation: Off

Alignment 2. Comparison of nucleotide sequence  
of SEQ ID NO:3 with the nucleotide sequence of  
Fig. 2 of Laux et al.

```

1 -----atgctatca 9
1 gcagtgtgtgaagattgtcacagctgctggttggagaaaacgggggtgggcggtgatca 60
      20      40      60
10 ggtaacgcaggagaa-ggag-caacagcctgcggaggttcggc----- 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 gggagaacaattccccggggacacctgcacgagaccctgggctctcaggaactccgccc 120
      80      100      120
51 -----cgccgcgggcccaggacctcatcagcgtcccc 81
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 aggtcttgccaattggggtgatcctgtagcgccgcgg====tttcagcatca===== 168
      140      160
82 cgcaacacctttatgacactgcttcagaccaacctggacaacaaaccgccgaggcagacc 141
=====

      160      180      200
142 ccgctaccctacgcggccccgctgcccccttttcccaccaggcaatagccaccgcgcct 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
169 =====caggttattttgcc= 182
      180
202 tcctacgggtcctggggc-----cg 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
183 t=gaagcttgctggggcgtaaatccctctcgcccttgtttctcagagagcatttcaggccg 241
      200      220      240
221 g-----a 222
   |
242 gttttgcagtcgctgctgcagctatggggtccctagaaatggtgccaatgggcgcgggtc 301
      260      280      300
223 gcgggtcgccccggcgccgggctactttacctccccaggaggttactacgccgggccccgcg 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 cccctagccccggc===== 315

      300      320
283 ggccggggaccgggtgccttcttggcgatggacgctcacacctac----- 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
316 ==ggggatccggatgggta====cgatggcggaacaactcccaatatccatctgct 366
      320      340      360
328 -----cacccccaccc----- 338
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 tctggctcttctgggaacacccccacccacccgaacgatgaggaacgtgaatctaataa 426
      380      400      420

```

```

-----
427 gagccccaccgccttatgaggaccatattggggcaatggcgaccgtcactcggactat 486
      .               .               .               .
      440               460               480
-----
487 caaccactaggaaccaagatcaaagtctgtacttgggattgcaacacgacgggaatgac 546
      .               .               .               .
      500               520               540
340
339 -----acacccccctccggcctac----- 357
      | ||||| |||
547 gggctccctcccccctcc==ctactctccacgggatgactcatctcaacacatatacgaa 603
      .               .               .               .
      560               580               600
-----
604 gaagcgggcagaggaagtatgaatccagtatgcctgcctgtaattggtgcgcctacctc 663
      .               .               .               .
      620               640               660
360
358 -tttggt-----tgccg----- 369
      ||||| |||
664 ttttggtggcggctattgccgcctcgtgtttcacggcctcagttagtaccgttgtgacc 723
      .               .               .               .
      680               700               720
-----
724 gccaccggttggccctctcacttctactcttggcagcagtggccagctcatatgccgct 783
      .               .               .               .
      740               760               780
-----
784 gcacaaaggaaactgctgacaccggtgacagtgttactgcggttgctactttctttgca 843
      .               .               .               .
      800               820               840
-----
844 atttgcctaacatggaggattgaggaccaccttttaattctcttctgtttgcattgctg 903
      .               .               .               .
      860               880               900
-----
904 gccgcagctggcggactacaaggcatttacgttctggtgatgcttgtgctcctgatacta 963
      .               .               .               .
      920               940               960
-----
964 gcgtacagaaggagatggcgccgtttgactgtttgtggcggcatcatgtttttggcatgt 1023
      .               .               .               .
      980               1000              1020
-----
1024 gtacttgtcctcatcgctgacgctgttttgcagctgagtcacctccttggagctgtaact 1083
      .               .               .               .
      1040              1060              1080
-----
1084 gtggtttccatgacgctgctgctactggctttcgtcctctggctctcttcgccagggggc 1143
      .               .               .               .
      1100              1120              1140
-----
1144 ctaggtactcttgggtgcagcccttttaacattggcagcagctctggcactgctagcgta 1203
      .               .               .               .
      1160              1180              1200

```

560 . 580 . 600 .  
557 gtgtgaatgaattacagcacacgctacaggccctgcgccgggaacgctgtctctacggcc 616

```

      620              640              660
617 acaccggagtcggatactgccccagcagggccctgctacaccactcggggccttacg 676
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1836 =====aaccacacggagcagggcaacat=tgcaggg=ctctgttcacg=cg 1878
      1840              1860
      680              700              720
677 gatttcagcctcatcaaagctacgaagtgccagatacgtccctca----- 722
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1879 atggtcgtcatc=tggct=ctcctgtgtgac=====ccctcactttgtacagactt 1927
      1880              1900              1920
723 -----tccgccccaccacca 738
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1928 ttggcaatgggagcacattccccccgcctttgggcaccccacggggtgctccccctggac 1987
      1940              1960              1980
      740              760
739 acttct-caccaggcagctca-----ggcgcagcctccac----- 772
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1988 acttatgtttcaagcagctcacctatgggtcactcaggcacggtcgccccctccgagtgacc 2047
      2000              2020              2040
773 -----ccccgg 778
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2048 agtcaccttccagactatgcatacactgaatttagcctgatattgtccccctagccccgg 2107
      2060              2080              2100
      780              800              820
779 gcacacaggccccccaagcccactgtgtggccgagtcacgatccctgaggcgggagcag 838
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2108 gc=c=cagccctcctcagaa===== 2125
      2120
      840              860              880
839 ccgggaactctg-----gaccccgaggagacaccaaccctcagcagcccacc 885
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2126 =====aactctgcatggagaagctggac===== 2148
      2140
      900              920              940
886 accgagggccaccaccgcggaaagaaactggtgcaggcctctgcgtccggagtggtcag 945
=====
      960              980              1000
946 tctaaggagcccaccaccccccaaggccaagtctgtgtcagcccacctcaagtccatcttt 1005
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2149 =====gtgaa=cct=ccccccagacct=gtgtgc=tgtat 2180
      2160              2180
      1020
1006 tgcgaggaattgctgaataaacgcg-tg-gcttga----- 1038
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2181 ttacaaacactac==aataaacccaatgtgcaaaaaaaaaaaaaaaaaaaaa 2227
      2200              2220

```

% Identity = 11.7 (323/2749)

///

9310-13DVCTDV SEQ ID NO 2.xprt x Laux et al. EBV terminal gene.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ\*

Protein sequence 498 aa MGSLEMVPMGAG ... ERPPTPYRNTV\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 3.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence encoded by the nucleotide sequence of Fig. 2 of Laux et al.

```

-----
1 MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTPPNDEERESNEEPPPPY 60
      .               20               .               40               .               60
-----
61 EDPYWGNGDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRGS 120
      .               80               .               100              .               120
-----
121 NPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALSLLLLAAVASSYAAAQRKLLT 180
      .               140              .               160              .               180
-----
181 PVTVLTAVVTTFFAICLTWRIEDPPFNSLLFALLAAAGGLQGIYVLVMLVLLILAYRRRWR 240
      .               200              .               220              .               240
-----
241 RLTVC GGIMFLACVLVLIVDAVLQLSPLLGA VTVVSM TLLLLAFVLWLSSPGGLGTLGAA 300
      .               260              .               280              .               300
-----
1 -----MARRLPKPTLQGRLEADFPD SPLL PKFQELNQN NLPN 37
      M + TL L S L K L + L
301 LLTLAAALALLASLILGTLNLT T MFLMLLWTLV VLLICSSC SSCP LSK=ILLARLFL=Y 358
      .               320              .               340              .
      40               60               80
38 DVFREAQRSYLVLFTSQFCYEEYVQRTFGVPRRQRAIDKRQRASVAGAGAHHLGGSSAT 97
      + S L+ S + + A + A GS
359 ALALLLLASALIAGGSILQTNFKSLSTEFIPNLF CMLLLIVAGILFILAILTEWGSGNR 418
      360               380               400
      100               120               140
98 PVQQAQAAASAGTGALASSAPSTAVAQSATPSVSSSISSLRAATSGATAAASAAA VDTG 157
      +A + T ++ + + + L A
419 TYGPVFMCLGGLLTMVAGAVWLT VMSNTLLSAWILTAGFLIFLIGFALFGVIRCCRYCCY 478
      420               440               460
      160
158 SGGGGQPHDTAPRGARKKQ* 177
      + + P R *
479 YCLTLESEERPPTPYRNTV* 498
      480

```

% Identity = 4.6 (23/500) % Homology = 3.2 (16/500) % Total = 7.8 (39/500)

///



9310-13DVCTDV SEQ ID NO 4.xprt x Laux et al. EBV terminal gene.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA\*

Protein sequence 498 aa MGSLEMVPMGAG ... ERPPTPYRNTV\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 4.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence encoded by the nucleotide sequence of Fig. 2 of Laux et al.

```

      20      40
1 MLSGNAGE-GATACG-GSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPPF 58
  M S      GA      G      G D + + +      N      P +      P PP+
1 MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTPPNDEERESNEEPPPPY 60
      20      40      60
60
59 SHQAIATAP--S-YGP-GAGAVAPAGGYFTSPG-GYYAGPAGG-DPGAFLAMD-AHTYHP 111
      S Y P G + G      G      P      D + + A
61 EDPYWGNGDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRGS 120
      80      100      120
112 HPHPPPAYFG--LPGLFGPPPPCLLTDSHLRADYVPAPSRSNKRKRDPDEEGGGLF- 168
  +P P      L L      C + S +      A S      L
121 NPVCLPVIVAPYFLWLAIAASCFTASVSTVVTATGLALSLLLLAAVASSYAAAQRKLLT 180
      140      160      180
169 PGEDATLYRKDIAGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQ 228
  P      T      A      E      L + G      Y      + ++
181 PVTVLTA VVTFFAICLTWRIEDPPFNSLLFALLAAAGGLQGIYVLVMLVLLILAYRRRW 240
      200      220      240
229 PHQSYEVPRIVPHPPPPPTSHQAAQAQPPPPGTQAPEAHCV-AESTIPEA-GAAGNSGPR 286
      ++      +      T      + A      + G G G
241 RLTVCGGIMFLACVLVLIVDAVLQLSPLLGA VTVVSM TLLLLAFVLWLSSPGGLGT LGAA 300
      260      280      300
287 EDTNPQQPTTEGHHRGKKLVQASASGVAQSKEPTTPKAKSV-SAHLKS-IFCEELLNKR 344
  T      L + +      S S+ S I L +
301 LLTIAAALALLASLILGTLNLTMTFLLMLLWTLVLLICSSSCSSCPLSKILLARLFYAL 360
      320      340      360
345 A*----- 346
  A
361 ALLLLASALIAGGSILQTNFKSLSTEFIPNLF CMLLLIVAGILFILAILTEWGS GNRTY 420
      380      400      420
-----
421 GPVFMCLGGLLTMVAGAVWLTVM SNTLLSAWILTAGFLIFLIGFALFGVIRCCRYCCYYC 480
      440      460      480
-----
481 LTLESEERPPPTPYRNTV* 498

```

% Identity = 10.0 (50/498) % Homology = 4.8 (24/498) % Total = 14.9 (74/498)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Laux et al. EBV terminal gene.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein sequence 498 aa MGSLEMVPMGAG ... ERPPTPYRNTV\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 5.** Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence encoded by the nucleotide sequence of Fig. 2 of Laux et al.

```

      20
1 AVDTGSGG-G-GQPHDTA-PRGARKKQ----- 24
      G G P      P G
1 MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTPPNDEERESNEEPPPPY 60
      20              40              60
-----
61 EDPYWGNDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRGSM 120
      80              100              120
-----
121 NPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALSLLLLAAVASSYAAAQRKLLT 180
      140              160              180
-----
181 PVTVLTAVVTTFFAICLTWRIEDPPFNSLLFALLAAAGGLQGIYVLVMLVLLILAYRRRWR 240
      200              220              240
-----
241 RLTVCGGIMFLACVLVLIVDAVLQLSPLLGAFTVVSMTLLLLAFVLWLSSPGGLGTLGAA 300
      260              280              300
-----
301 LLTLAAALALLASLIIGTLNLTMTFLLMLLWTLVLLICSSSCSSCPLSKILLARLFLYAL 360
      320              340              360
-----
361 ALLLLASALIAGGSILQTNFKSLSSSTEFIPNLFMLLLVAGILFILAILTEWGSNGRTY 420
      380              400              420
-----
421 GPVFMCLGGLLTMVAGAVWLTVMSNTLLSAWILTAGFLIFLIGFALFGVIRCCRYCCYYC 480
      440              460              480
-----
481 LTLESEERPPTPYRNTV* 498

```

% Identity = 1.0 (5/498)      % Homology = 0.0 (0/498)      % Total = 1.0 (5/498)

///

9310-13DVCTDV SEQ ID NO 6.xprt x Laux et al. EBV terminal gene.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 498 aa MGSLEMVPMGAG ... ERPPTPYRNTV\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

Alignment 6. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence encoded by the nucleotide sequence of Fig. 2 of Laux et al.

```

-----
1 MGSLEMVPMGAGPPSPGGDPDGYDGGNNSQYPSASGSSGNTPTTPNDEERESNEEPPPPY 60
      .                20                .                40                .                60
-----
61 EDPYWGNDRHSDYQPLGTQDQSLYLGLQHDGNDGLPPPPYSPRDDSSQHIYEEAGRGS 120
      .                80                .                100               .                120
-----
121 NPVCLPVIVAPYLFWLAAIAASCFTASVSTVVTATGLALSLLLLAAVASSYAAAQRKLLT 180
      .                140               .                160               .                180
-----
181 PVTVLTAVVTTFFAICLTWRIEDPPFNLLFALLAAAGGLQGIYVLVMLVLLILAYRRRWR 240
      .                200               .                220               .                240
-----
241 RLTVC GGIMFLACVLVLIVDAVLQLSPLLGA VTVVSM TLLLLAFVLWLSSPGGLGTLGAA 300
      .                260               .                280               .                300
-----
301 LLTLAAALALLASLILGTNLNLTMTFLLMLLWTLVVLICSSCSCPLSKILLARLFLYAL 360
      .                320               .                340               .                360
-----
361 ALLLLASALIAGGSILQTNFKSLSTEFIPNLF CMLLLIVAGILFILAILTEWGS GNRTY 420
      .                380               .                400               .                420
-----
1 -----STAVAQSATPSV 12
      +
421 GPVFMCLGGLLTMVAGAVWLTVMSNTLLSAWILTAGFLIFLIGFALFGVIRCCRYCCYYC 480
      .                440               .                460               .                480
      20
13 SSSISSLRAATSGATAAA 30
      + S R T
481 LTLESEERPPTPYRNTV* 498

```

% Identity = 0.6 (3/498)      % Homology = 0.4 (2/498)      % Total = 1.0 (5/498)

///

9310-13DVCTDV SEQ ID NO 1.xdna x Bankier et al. EcoRI Dhet fragment.xdna => DNA Parallel

DNA sequence 538 bp catgatggcacg ... aaacagtagccc linear

DNA sequence 12436 bp gaattctcaaag ... tgtttagaattc linear

Method: Blocks (Martinez)  
Layout: Standard  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Translation: Off

Alignment 7. Comparison of nucleotide sequence of SEQ ID NO:1 with the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1  gaattctcaaaggcggcaccctcgccggcgcgccctgtcctcccagggacccgagacgaag 60
    .                20                .                40                .                60
-----
61  gcccgctctgtagaggaagtgggttgcgcattgcggggccagctcccagtagaccacgtccccc 120
    .                80                .                100               .                120
-----
121 cagacgcgcaggcacagggtctcggtcagggtctcgctctgttgccaggcaggactgc 180
    .                140               .                160               .                180
-----
181 agcttggccagaccctcggtggccacctggcgcaggtaactgctccttgcgcttgagcgcg 240
    .                200               .                220               .                240
-----
241 tccgagagggcgccggacggggccgggctctcgtgccccagccggccggggcacctccggg 300
    .                260               .                280               .                300
-----
301 ctctcccggaagcctcctcctcgccctcgccccaaccgctgcatggctcggttgagccgc 360
    .                320               .                340               .                360
-----
361 gtgtacagctcgttcctcttttgcaggatggcccggtactgggggtgcgccgtgaaggcg 420
    .                380               .                400               .                420
-----
421 gcggcgagtcgcgccttcagcgcctccaccgctcgcccgaggagctgtagaccccgccg 480
    .                440               .                460               .                480
-----
481 cagaagagccgctccgtggccccgggagccacggcgctcaaacaggtgagtcagccttgcc 540
    .                500               .                520               .                540
1  -----catg 4
    |
541 cccgccagcgcctcctcgccaggccccccgcaccagggccaggcgacgctcccgggcaaac 600
    .                560               .                580               .                600

```

```

      20
5  atggcagcgccggctgcccaag-c-ccaccctcc----- 35
   | | | | | | | | | | | | | | | | | | | | |
601 agggcagagagggcggaatggccgccaccctccccctgccccgttgccacgatagcatgg 660
      620      640      660
-----
661 ccgccagagttccaatagaggagctccgagagctccgccacctccgggggcactgtcgag 720
      680      700      720
-----
721 aagacgttgtaggtgtccagcgctctggctgccccctctgcctccggccgccccggggccc 780
      740      760      780
-----
781 gggaccgcgcctctctctggggccgcccggcctcgcttctctcagcctccaacaggtgc 840
      800      820      840
-----
841 ccgagccccgcctggcgggacttcattctcaaacagtcccgagaccggctccggattcacc 900
      860      880      900
-----
901 ggcaccgccaggtggttacaggagacgtgggtccccctctgccgtggaagggttgccgtgg 960
      920      940      960
-----
961 ttgggcagaaccatcagctcgcccacacagcgccagcagggcacagaggtgatgtagagg 1020
      980      1000      1020
-----
1021 cgcgggtctgggatgggacttacgccccgaaagcggcccagcagatccaggggcccgttcc 1080
      1040      1060      1080
-----
1081 aggctctccagccccatgggtgtgagacatgcaataaaacacgctattgattctcttcatt 1140
      1100      1120      1140
-----
1141 aaaatctctatgtcattttattaggcacaaacttacatcgactttatgccccccgtaaaac 1200
      1160      1180      1200
36 -----a 36
1201 tccacagagtacgcgactgagggggtacggagaggcgggacccgggtaccctttctacca 1260
      1220      1240      1260
      40
37 ggggag----- 42
   | | | | |
1261 ggggcgagcagcgcggcagaggcctctctcgagttctctagcaggtgcaccagctccagg 1320
      1280      1300      1320
-----
1321 gacagggcgctgcatgcacggtcattctgccgtctcaaacggggaaggaggatggcctcc 1380
      1340      1360      1380

```

1381 agctcgggccagcaggccggcggttgcgccaccgcagccacgtccagactccgggggtcc 1440  
 . 1400 . 1420 . 1440

1441 agccgggtgcacacgctcagctcaaccgccaggcggtacacctggctgtacgccgcgcgc 1500  
 . 1460 . 1480 . 1500

1501 agcagccccgacatcgccgccccaggggtctctagacctcgagtcgggggagaacggtgg 1560  
 . 1520 . 1540 . 1560

1561 ccagacggcgcttgcggtctgccccggagccctgccctcctccaccagcagcagcccg 1620  
 . 1580 . 1600 . 1620

1621 ccgaggcctgcgcgcgggtgctgaccggctcggccacgctgataaagtgtctctgggctg 1680  
 . 1640 . 1660 . 1680

1681 ccccgggcccccacactccctccagaaagtcccgagcggcctccgcggtccactcta 1740  
 . 1700 . 1720 . 1740

43 ---gctggaggc----- 51  
 |||||  
 1741 tcccgctggaggcaatggtcgccagggtttctaggacgtgtccgccaggacggagaagc 1800  
 . 1760 . 1780 . 1800

1801 ggccaataagtactccgcgtcgccctagtcagcgaggcgcatgcctcgcccatggcat 1860  
 . 1820 . 1840 . 1860

1861 ccacaagggtgcacaccacatcaaacacacagtgcttctctctgtttttgtgatataatgg 1920  
 . 1880 . 1900 . 1920

1921 cctccaggccagccctgatgttctcaatctcatatgtggtcgcggcttgggtcgggcgct 1980  
 . 1940 . 1960 . 1980

1981 tcacggtcaaccctaggggtgggggtggcaagacaaacttcttccgcatggaagagcccc 2040  
 . 2000 . 2020 . 2040

[illegible]

2101 ccataaaagtatcccaggttcccggcctggaatatctggttggtgcggttgacccccgtgt 2160  
 . 2120 . 2140 . 2160

```

52 -----ggattttc-cagac 64
2161 acttgttgatgggtcactggcagcgtgacaaccggacgggccttgacagacctggctaagac 2220
      2180      2200      2220
65 agtccccctgcttcctaaatttcaagagctgaaccagaataatctccccaatgatgttttt 124
      80      100      120
2221 agtc=====tgtggccgcgcag=accaccgt==ggt=cgcagt=aagggagg 2262
      2240      2260
125 cgggagggtcaaagaagttacctggtattttctgacatcccagttctgctacgaagagtac 184
      140      160      180
2263 aggtggcctccgcgtag==gcc==g==ctgccgac=tccaccgcccgc=gtgcccagtag 2314
      2280      2300
185 gtgcagaggacttttggggtg-----cctcggc-----gccaacg 219
      200
2315 gtgggg=gtagtcacgggcgggcaccgactgcgtcctcggcaccagtcacctgaatcaggc 2373
      2320      2340      2360
220 cgccatagacaagaggcagagagccagtggtggctggggctggtgctcatgcacaccttg 279
      240      260
2374 tgatgtagaactgggtctggccgcacgccttcaggatggcggttggtgagcctctgcttg 2433
      2380      2400      2420
280
280 cg-----ggtcatccgcc 292
      280
2434 cgtaagtaccaggttgccaggcaccacatctatgacgttgctctcttcgtgggcccggg 2493
      2440      2460      2480
300
293 a-ccccgtcca----- 303
      300
2494 agcccccggtccacaaagagggccaggtcagagtactcctccgcgctggccccgctgggga 2553
      2500      2520      2540
-----
2554 cagggaccgagcgccgcctggaaaagttgtgccacaggtacaggcttgagagcttagtgt 2613
      2560      2580      2600
-----
2614 ccgggaatagggtccttggttaggtggtgaggaatttcagtagggcccgttgatgatgt 2673
      2620      2640      2660
-----
2674 agttctccctcctggttagtgacttgatgaagctgttctggaggcggcattctcccccg 2733
      2680      2700      2720
-----
2734 tgaagaccacctgttcttgatcttgatgttcctggggcacagcatcagcaccttgga 2793
      2740      2760      2780
-----
2794 tgcgcacaggcagccgccggccgtacaccggccctgcaggcgccgctccaggtctggca 2853
      2800      2820      2840
-----
2854 ggtcgcaggtgggctccccatgcaccaccttgccctccttgcccgtaggaccccccttgt 2913
      2860      2880      2900

```

```

304 -----gcaggctc----- 311
      |||||
2914 cgatggccaggctcctaagttggtgcacagcgtcttggtagtgaccctttagccactctg 2973
      2920                2940                2960

-----

2974 gggggctctggccaagcccggttgctcattctcatagcacatacagatgggcagggaga 3033
      2980                3000                3020

-----

3034 tgtcctgcaggatggtcagcagtgagcggtaaaacagctgggtgaagatggggcaggcgg 3093
      3040                3060                3080

-----

3094 gctgcgcaaaggggttgacagagtactgcatcacgtggtagcagctcttgaccaggctcct 3153
      3100                3120                3140

312 -----aggccg 317
      |||
3154 tgtaggatgatgttgttcttgccatgctgttcataaaactggaccacttcggcgtccaccg 3213
      3160                3180                3200
      320
318 ccgcatcc----- 325
      |||||
3214 ccgcatccacgtccttgaacatcttgacaaaagtcacgcggggccatggggctccttctcta 3273
      3220                3240                3260

-----

3274 gctttccttcagcgtctatgcccagccgagacagccgctccagcaggttctggttcagct 3333
      3280                3300                3320

-----

3334 gccagtaggtgtagcggggctcgtcgtccggccgctgcccgctcctccttatcgatga 3393
      3340                3360                3380

-----

3394 agttgagaaaagttgcccaaaaagtcctgtctcgttgtaggagcccagggcccccagatatca 3453
      3400                3420                3440

326 -----gctgg----- 330
      ||||x
3454 catagggggtccctccgctgcgtggacatgacgggggggaagcgggtccctcagcctaaaga 3513
      3460                3480                3500

-----

3514 agagcgtgttcaggcacacggccggggcccgccctcgcagagcgagcacatgggactgg 3573
      3520                3540                3560

-----

3574 cggccgcccccgccacgtagctgcccgtctccggccaccggggtcagagagctcttctgtc 3633
      3580                3600                3620

-----

3634 cctggcaaaaactgcaggtagtaggcatacggggcaagaaggttgggcgagaaggaggccg 3693
      3640                3660                3680

```



-----  
 3694 catagaccaggtgctccacagcgtagtttcccggaccggtgggttccgggtcacgtctggcc 3753  
       3700                              3720                              3740  
 -----

3754 caccacgcccagagaagcaggggtcggcggcaggggtcccagggtcccctcctgcaggggtcc 3813  
       3760                              3780                              3800  
 -----

3814 ccaggccgtgggtcatgtagaaactgttaaagagactctccttgccctgaccggttgact 3873  
       3820                              3840                              3860  
 -----

3874 tcgagacccccgagacgtagaggacggaattgggtggcaaagatctgcgtggacacgtggg 3933  
       3880                              3900                              3920  
 -----

3934 gggccagggtggcattatatcgggtgtaacgcagccacacgggcctctggaccctcacagt 3993  
       3940                              3960                              3980  
 -----

3994 cggcaaacagggggccacgagtcgtagttgagggtggccgggggtctcgtgcgaggcctcca 4053  
       4000                              4020                              4040  
 -----

4054 gcattggcgggtgcgtagctcaccgccagctcgcattgccgcgctgtccacaatcattaagg 4113  
       4060                              4080                              4100  
 -----

4114 ctcccaggtccgggtgactgatgggtgagggtgggaactccttgaggggggcccaccttgg 4173  
       4120                              4140                              4160  
 -----

4174 ccaccttggcctgggtcctgcaggctctgcttctccagcagctccaccagcttgcccaccc 4233  
       4180                              4200                              4220  
 -----

4234 gtccgacgcgcagcgcctgcgccagcccgggtgtacagcgcctcgtgcattgcagcggctga 4293  
       4240                              4260                              4280  
 -----

4294 ggtccgagttgtaaaactggcggagctggggcacgccctctgggaacacctccttgctgt 4353  
       4300                              4320                              4340  
 -----

4354 agagcgggaccctaacgctcgcagactgccccaccgctacctcctgttttaacgatggaa 4413  
       4360                              4380                              4400  
 -----

4414 tggccaccaggtttccgctgtagagtcgctccttgaaggcctcggttattgccaccgccc 4473  
       4420                              4440                              4460  
 -----

4474 ccaggtaggcagaggggatctagcccttcggggaagaagtccccggcttgagctttccc 4533  
4480 . 4500 . 4520 .

4534 tcggtagggcgctgtaggcgctgtacccaaacacctccctgggtctcgccacagagggcct 4593  
4540 . 4560 . 4580 .

4594 cgagaccgcggcccoctcaaagatggggggaaccatatgggcattgtggaacacgtagatgt 4653  
4600 . 4620 . 4640 .

4654 ccctgtgataggaggttagcgcgtaggagcccgagttggggtcgggcctcctgtgcagag 4713  
4660 . 4680 . 4700 ..

4714 ccttgacattgatgctgaagcccggtccacggtgatgccgcaaaggagcggcaccgtca 4773  
4720 . 4740 . 4760 .

4774 ggcacctgtggcccgcgtagccgggtccccagtggtggccacctccctaagagggtaggtgg 4833  
4780 . 4800 . 4820 .

4834 ccagggggtaaaagtagatgtagccgcacggaccgggctggctctggctgccagattat 4893  
4840 . 4860 . 4880 .

4894 cctcgctagctctgtgcaccctgcatgatgcccaagggtatcgccccggcctcccagttccca 4953  
4900 . 4920 . 4940 .

4954 cattaaatgttacactttactcatcacgcaacaccacctgtttattcatttacaagatt 5013  
          4960                4980                       5000

5014 tcaggaagtcagtcaggctggccagggccacgtcacggggaactgacgtctcagcgatc 5073  
5020 . 5040 . 5060 .

5074 ttggcatgccgcccagcctcgcaaacagagtctgcgatagagggccaggtagtgggcga 5133  
5080 . 5100 . 5120 .

5134 ttgccccagcacgaaggcggcgtctttgtgggtcatccaggtagtttcgcaccgcaaaca 5193  
5140 . 5160 . 5180 .

5194 ccactgtgtagcacagcaccaccctgagccgcgaccagtagtcgtagtggtcgttgtaca 5253  
5200 . 5220 . 5240 .

-----  
 5254 ctgcgcgcaggacgctgatgatgagccgtacgtgcgtgtctttgccccgatgtcggctg 5313  
       5260                                  5280                                  5300  
 -----

5314 tcctgcaggccagctccgcgtacagcttccctatcccttcctcaggaggccttgatgagcc 5373  
       5320                                  5340                                  5360  
 -----

5374 ggcagaggaccagggctggcaaaggcaggtctttctcatcccggtgaacaccgcgtaca 5433  
       5380                                  5400                                  5420  
 -----

5434 tggccctgaacatgaggtagctggactcagccaccttgctcgtccggcgaggcgcgga 5493  
       5440                                  5460                                  5480  
 -----

331 -----gaccgggg----- 338  
           |||||  
 5494 cccacgcctcgaccggggctcctcacaacacagaatctgtagacttggtggcctcatgg 5553  
       5500                                  5520                                  5540  
 -----

5554 tctcgtcaggccagctcacgggcttcaggcttatatgataaaatgggcgtggcagaatag 5613  
       5560                                  5580                                  5600  
 -----

5614 tataagacgcgaggcctgggtgaggagagtccagagcaatggccaggttcacgctcagc 5673  
       5620                                  5640                                  5660  
 -----

5674 tcctcctgttgccctcctgtgtggccgccggccaggctgtcacccgctttcttgggtgagc 5733  
       5680                                  5700                                  5720  
 -----

5734 gagtcaccctgacctcctactggaggaggggtgagcctcgggtccagagattgaggtcagct 5793  
       5740                                  5760                                  5780  
 -----

5794 gggtttaaactgggcccaggagaggagcaggtgcttattgggcgcacatgcaccacgatgtca 5853  
       5800                                  5820                                  5840  
 -----

5854 tctttatagagtggcctttcaggggcttctttgatatccacagaagtgccaacaccttct 5913  
       5860                                  5880                                  5900  
 -----

5914 ttttagtagtcaccgctgccaacatctcccatgacggcaactacctgtgccgcatgaaac 5973  
       5920                                  5940                                  5960  
 -----

5974 tgggcgagaccgaggtcaccaagcaggaacacctgagcgtggtgaagcctctaacgctgt 6033  
       5980                                  6000                                  6020  
 -----

-----  
 6034 ctgtccactccgaaagggtctcagttcccagacttctctgtccttactgtgacatgcaccg 6093  
           6040                                  6060                                  6080

-----  
 6094 tgaatgcatttccccatccccacgtccagtggtcatgcccagggcggtggagcccgcac 6153  
           6100                                  6120                                  6140

-----  
 6154 caactgctggcaaatggcggtgttatgaaggaaaaggatgggagcctctctgttgctgttg 6213  
           6160                                  6180                                  6200

-----  
 6214 acctgtcacttcccaagccctggcacctgccagtgacctgcgttgggaaaaatgacaagg 6273  
           6220                                  6240                                  6260

-----  
 6274 aggaagcccacggggtttatgtttctggatacttgtcgcaataaacgcacttgcctat 6333  
           6280                                  6300                                  6320

-----  
 6334 caccttgtttttagtgtggcattgggggggtggcattgcgggtggatagcctcgcgactcg 6393  
           6340                                  6360                                  6380

-----  
 6394 tgggaaaaatggcggaaggggcacctggtgggaaaaatagttccaggtgacagcagcagtgtgt 6453  
           6400                                  6420                                  6440

-----  
 6454 gaagattgtcacagctgctggtttggagaaaaacgggggtggcggtgatcaggggagaaca 6513  
           6460                                  6480                                  6500

-----  
 6514 attccccggggacacctgcacgagaccctgggctctcaggaactccgcccagggtcttgc 6573  
           6520                                  6540                                  6560

-----  
 6574 caattggggtgatcctgtagcgccgcggtttcagcatcacagggttattttgcctgaagct 6633  
           6580                                  6600                                  6620

-----  
 6634 tgctggggcgtaaatccctctcgccttgtttctcagagagcatttcaggccggttttgca 6693  
           6640                                  6660                                  6680

-----  
 6694 gtcgctgctgcagctatggggtccctagaaatggtgccaatggcgcggtccccctagc 6753  
           6700                                  6720                                  6740

-----  
 6754 cccggcggggatccggatgggtacgatggcggaacaactcccaatatccatctgcttct 6813  
           6760                                  6780                                  6800

-----  
6814 ggctcttcttgggaacacccccacccacccaacgatgaggaacgtgaatcctaagaagag 6873  
6820 . 6840 . 6860 .

-----  
6874 cccccaccgccttatgaggacccatattggggcaatggcgaccgtcactcggactatcaa 6933  
6880 . 6900 . 6920 .

-----  
6934 ccactaggaacccaagatcaaagtctgtacttgggattgcaacacgacgggaatgacggg 6993  
6940 . 6960 . 6980 .

-----  
6994 ctccctccccctccctactctccacgggatgactcatctcaacacatatagaagaagcg 7053  
7000 . 7020 . 7040 .

-----  
7054 ggcagaggaaggtaagagtgccatctatctgtactttttatttattgcatcacaaagtcaca 7113  
7060 . 7080 . 7100 .

-----  
7114 tcaataataagggcgccatctagcgggagatgttatccacaccatcccaattcacatctc 7173  
7120 . 7140 . 7160 .

-----  
7174 agggacaacagggtcaaagttctttgttgacacccccagcgctggctccaggggggtggaag 7233  
7180 . 7200 . 7220 .

-----  
7234 cgttggatgcagtcctccgcacgccccggcgacgcctcctcccaacgcgtttctgaggatc 7293  
7240 . 7260 . 7280 .

-----  
7294 agtcgctggctggtgggcatcggagtcggtggcggtcctccacggggacacgctccttc 7353  
7300 . 7320 . 7340 .

-----  
7354 ttggccttggtcttttgaccttttgacattcttctgaaggaacggcgagagtagcgtag 7413  
7360 . 7380 . 7400 .

-----  
7414 aatccagccagtggtctacccggtcgcatggtggcttcttagatgaggagcaggcataaa 7473  
7420 . 7440 . 7460 .

-----  
7474 agtccaaacaggacacagagtaccaccaggagtagtcttagtctgctgacgtctgggtcc 7533  
7480 . 7500 . 7520 .

-----  
7534 tcggggcaggggtggctaggcctggtctccgtagaagagccgggcaggccgcaggcagag 7593  
7540 . 7560 . 7580 .

---

-----

-----

-----

-----

340

-----

---

-----

.....

-----

360

```

342 -----tggcatcatcagcgcgcgtcca-cggccg 368
      | | | | | | | | | | | | | | | | | |
8314 atagtagcttagctgaactgggccgtgggggtcgtcatcatc=====tccaccggaac 8366
      8320                      8340                      8360

```

```

      380              400              420
369 tagcccagtcgcgcgacccccctctgtttcttcatctattagcagcctccgggcccgcgactt 428
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8367 cagaagaacccaaaagcagcgttaggaaggt=gtggatca=ccgccgccatggc=cggaat 8423
      8380              8400              8420
      440              460
429 cggggggcgactgccgcgcctccgcc-gccgcagccgtc----- 466
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8424 c=atgactatgaccgccgcctccgtctgtcatcaaaggcgggccctggtcacctcctttg 8482
      8440              8460              8480
-----
8483 ttttcaacctcttccgtcaattgtggagggcctccatcatttccagcagagtcgctaggg 8542
      8500              8520              8540
-----
8543 ctatgaggcagcgggtcatgtgggccattgtcatcagtggtgtcagggtcctgtggggcca 8602
      8560              8580              8600
-----
8603 ttgtcatcagtggtgtcagggtcctgaggcagcgggtcatgtgggccattgtcatcagtg 8662
      8620              8640              8660
-----
8663 ttgtcagggtcctgtggggccattgtcatcagtggtgtcagggtcctgtggggccattgtca 8722
      8680              8700              8720
-----
8723 ggaccacctccaggtgcgcctaggttttgagagcagagtggggggtccgtcgccgggtcca 8782
      8740              8760              8780
-----
8783 ctcacgagcaggtggtgtctgcccctcgttgaggttagagtcagattcatggccagaatca 8842
      8800              8820              8840
-----
8843 tcggtagcttggttgaggggtgcgggagggagtcacgtgggtggtgttcactgtgtcgt 8902
      8860              8880              8900
-----
8903 tgtccatggtaatacatccagattaaaatcgccagaaacaggaggagccaaaggagatca 8962
      8920              8940              8960
-----
8963 accaatagagtccaccagttttgtttagatagagagcaataatgagcaggatgaggtct 9022
      8980              9000              9020
-----
9023 aggaagaaggctaggaagaaggccaaaagctgccagatggtggcaccaagtgcgcagagc 9082
      9040              9060              9080
467 -----gatacc----- 472
      |||||
9083 atctccaataagtagatccagatacctaagactgcgttgaaaaaagagtgttaggggttg 9142
      9100              9120              9140

```

```

9143 aaaagtgggggtgtggtaaataattcctaggggaatgttagatcttaccaagtaagcacc 9202
      .               9160               .               9180               .               9200
      -----

9203 gaagatgaacagcacaaattccaaggaacaatgcctgtccgtgcaaattccagagagcgat 9262
      .               9220               .               9240               .               9260
      -----

9263 gagcaggaggggtgactggggaaagaggagaaagtgcgttagagaaggaagagtaagggaa 9322
      .               9280               .               9300               .               9320
      -----

9323 aggggggtgtggggc aaaggggtgtaatacttactcatcagtaggagtatacaaagggctcc 9382
      .               9340               .               9360               .               9380
      -----

9383 aagtggacagagaaggtctcttctgaagataaagatgatcaaaattataattataagcat 9442
      .               9400               .               9420               .               9440
      480               500               520
473 -----gggt-caggtggcgggggacaacccacgaca-cgcgccacgcgggggca 520
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
9443 gagagcaaaggaatagaggacaaggagggtcctccagtcagtcactcataacgatgta 9502
      .               9460               .               9480               .               9500
      -----

521 cgtaagaaacagtagccc----- 538
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
9503 cagccaaaacagtagcgccaagaggaggagaaggagagcaaggcctaggggaagaggagag 9562
      .               9520               .               9540               .               9560
      -----

9563 ggggggtcctcgagggggccgctcgcgggcccgggtgggcccctctcaaggtcgtgttccat 9622
      .               9580               .               9600               .               9620
      -----

9623 cctcagggcagtggtgtcaggagcaaggcagttgaggaaagaagggggcagagcagtggtga 9682
      .               9640               .               9660               .               9680
      -----

9683 gaggccttatgtagggcggtacgtcagagtaacgcgtgtttcttgggatgtaggccccggg 9742
      .               9700               .               9720               .               9740
      -----

9743 gggatttgcggggtctgcccggaggcaggtacgggtacagatttcccgaagcggcggtgtg 9802
      .               9760               .               9780               .               9800
      -----

9803 tgtgtgcatgtaagcgtagaaaaggggaagtagaaagcgtgtgtttgtgttagaaaagcgg 9862
      .               9820               .               9840               .               9860
      -----

9863 gtccccggggggc aaagctgtgggaatgcgggtggccaagtgcaacaggaaatggaaaggca 9922
      .               9880               .               9900               .               9920

```



-----

-----

-----

-----

1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2423
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

-----

.....

.....

.....

.....

-----

.....

10643 ggccccagggtgaccccggaacgtgaccggggctccggggtgacccagccaagcgt 10702  
 . 10660 . 10680 . 10700

10703 gaccaaggggcccgtgggtgacacaggcaaccctgacaaagggccccccaggaaagacccc 10762  
 . 10720 . 10740 . 10760

10763 cgggggggcatcgggggggtggggcatggggggccgcgcattcctggaaaaagtggaggggg 10822  
 . 10780 . 10800 . 10820

10823 cgtggccttccccgcgagcccccagcccccgcacagagcggcgctacggcgggcggg 10882  
          '                10840               '                10860               '                10880

10883 cggcggggggtcggggtccgcggggctccgggggctgcgggcggtggatggcggcggacgt 10942  
          .                .                .

10943 tccgggggatcgggggggtcggggggcgccgcgaggccatgcgtgaccgtgatga 11002  
          '                10960               '                10980               '                11000

11003 gggggcagggtcgcaggggggtgtgtctggtgggggcgggagcggggggcggcgcgggagc 11062  
 . 11020 . 11040 . 11060

11063 ctgcacgccgttgagggtagaatgacagggggcggggacagagaggcggtcgcgcctccc 11122  
 . 11080 . 11100 . 11120

11123 ggccgcgccagccaagcccccaagggggggcggggagcgggcaatggagcgtgacgaagg 11182  
 . 11140 . 11160 . 11180

11183 cccaggggtgaccccggaacacgtgacccgggggtccgggggtgaccagccaagcgtga 11242  
 . 11200 . 11220 . 11240 .

11243 ccaaggggcccgtgggtgacacaggcaacctgacaaagggccccccaggaaagacccccg 11302  
          '                11260               '                11280               '                11300

11303 tggggcatggggggccgcgcattcctggaaaaagtggagggggcggtggccttcccccgcg 11362  
 . 11320 . 11340 . 11360

11363 gccccccagccccccgcacagagcggcgctacggcgggcgggcggcggggggtcgggggt 11422  
 . 11380 . 11400 . 11420

11423 ccgcgggctccgggggctgcgggcggtggatggcggcggacgttccggggatcggggggg 11482  
 . 11440 . 11460 . 11480

-----  
11483 tcggggggcgccgcgcggggcgagccatgcgtgaccgtgatgagggggcagggtcgcagg 11542  
. 11500 . 11520 . 11540  
-----

11543 ggggtgtgtctggtggggggcgggagcggggggcgccgcgggagcctgcacgccgttgagg 11602  
. 11560 . 11580 . 11600  
-----

11603 gtagaatgacagggggcggggacagagaggcggtcgcgcccccgccgcgccagccaagc 11662  
. 11620 . 11640 . 11660  
-----

11663 cccaagggggcggggagcgggcaatggagcgtgacgaaggggccccagggctgaccccg 11722  
. 11680 . 11700 . 11720  
-----

11723 gcaaacgtgacccggggctccggggtgacccagccaagcgtgaccaagggggcccggtgggt 11782  
. 11740 . 11760 . 11780  
-----

11783 gacacaggcaaccctgacaaaggccccccaggaaagacccccggggggcatcgggggggtg 11842  
. 11800 . 11820 . 11840  
-----

11843 gggcatggggggcgccgcattcctggaaaaagtggagggggcggtggccttcccccgggc 11902  
. 11860 . 11880 . 11900  
-----

11903 cccccagcccccccgacagagcggcgctacggcgggcgggcggggggtcgggggtcc 11962  
. 11920 . 11940 . 11960  
-----

11963 gcgggctccgggggctgcgggcgggtggatggcgggcgacgttccggggatcgggggggtc 12022  
. 11980 . 12000 . 12020  
-----

12023 ggggggcgcgcgcggggcgagccatgcgtgaccgtgatgagggggcagggtcgcagggg 12082  
. 12040 . 12060 . 12080  
-----

12083 gtgtgtctggtggggggcgggagcggggggcgccgcgggagcctgcacgccgttgagggt 12142  
. 12100 . 12120 . 12140  
-----

12143 agaatgacagggggcggggacagagaggcggtcgcgcccccgccgcgccagccaagccc 12202  
. 12160 . 12180 . 12200  
-----

12203 ccaagggggcggggagcgggcaatggagcgtgacgaaggggccccagggctgaccccggc 12262  
. 12220 . 12240 . 12260

12263 aaacgtgacccggggctccggggtgaccagccaagcgtgaccaaggggcccggtgggtga 12322  
 . 12280 . 12300 . 12320

12323 cacaggcaaccctgacaaaggccccccaggaaagacccccggggggcatcgggggggtg 12382  
 . 12340 . 12360 . 12380

12383 ttggcgggggcatgggggggtcggatttcgcccttattgccctgtttagaattc 12436  
          .                .                .

```
% Identity = 2.3 (291/12474)
```

///

9310-13DVCTDV SEQ ID NO 1.xdna x Bankier et al. EcoRI Dhet fragment complement.xdna => DNA Parallel

DNA sequence 538 bp catgatggcacg ... aaacagtagccc linear

DNA sequence 12436 bp gaattctaaaca ... ctttgagaattc linear

Method: Blocks (Martinez)  
Layout: Standard  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Translation: Off

Alignment 8. Comparison of nucleotide sequence of SEQ ID NO:1 with the complement of the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1  gaattctaaacagggcaataagggcgaaatccgaccccccatgccccgccaacacccc 60
    .                20                .                40                .                60
-----
61  ccccgatgcccccgggggtctttcctggggggcctttgtcagggttgctgtgtcaccc 120
    .                80                .                100               .                120
-----
121 acgggcccccttggtcacgcttggtcgggtcacccccggagccccgggtcacgtttgccggg 180
    .                140               .                160               .                180
-----
181 gtcagccctggggcccttcgtcacgctccattgcccgtccccgcccccccttgggggctt 240
    .                200               .                220               .                240
-----
241 ggctggcgcgggccggggggcgcgaccgcctctctgtccccgccccctgtcattctaccctc 300
    .                260               .                280               .                300
-----
301 caacggcgtgcagggtccccgcgcgcccccccgctcccgccccaccagacacacccccctg 360
    .                320               .                340               .                360
-----
361 cgaccctgccccctcatcacggtcacgcatgggtgcgccccgcgcgcgccccccgacccc 420
    .                380               .                400               .                420
-----
421 cccgatccccggaacgtccgcccgcctccaccgccccgcagcccccgagcccgcggaccc 480
    .                440               .                460               .                480
-----
481 cgaccccccgccgccccgccccgcccgtagcgcgctctgtcgggggggctggggggcccgcg 540
    .                500               .                520               .                540
-----
541 ggggaaggccacgccccctccactttttccaggaatgcgcggccccccatgccccacccc 600
    .                560               .                580               .                600

```

```

-----
601 ccgatgccccccgggggtcttttcttggggggcctttgtcagggttgcctgtgtcacccac 660
    .               620             .               640             .               660
-----
661 gggccccttggtcacgcttggtcggtcaccccgaggccccgggtcacgtttgccgggggt 720
    .               680             .               700             .               720
-----
721 cagccctggggcccttcgtcacgctccattgcccgctccccgcccccttgggggcttgg 780
    .               740             .               760             .               780
-----
781 ctggcgcgggccggggggcgcgaccgcctctctgtccccgccccctgtcattctacctcca 840
    .               800             .               820             .               840
-----
841 acggcgtgcaggctccccgcgcgcgcccccgctcccgccccaccagacacacccccctgcg 900
    .               860             .               880             .               900
-----
901 accctgccccctcatcacgggtcacgcatggctgcgcgcgcggcgccccccgaccccccc 960
    .               920             .               940             .               960
-----
961 cgatccccggaacgtccgcgcgccatccaccgcccgcagcccccgaggcccgcggaaccccg 1020
    .               980             .              1000             .              1020
-----
1021 accccccgcccgcgccgcgcgcgtagcgccgctctgtgcgggggggctggggggccgcggg 1080
    .              1040             .              1060             .              1080
-----
1081 ggaaggccacgccccctccactttttccaggaatgcgcggccccccatgccccacggggg 1140
    .              1100             .              1120             .              1140
-----
1141 tttttcttggggggcctttgtcagggttgcctgtgtcacccacggggcccccttggtcacgc 1200
    .              1160             .              1180             .              1200
-----
1201 ttggtcgggtcaccccgaggccccgggtcacgtttgccgggggtcagccctggggcccttc 1260
    .              1220             .              1240             .              1260
-----
1261 gtcacgctccattgcccgctccccgcccccccttgggggcttggctggcgcgggccgggggc 1320
    .              1280             .              1300             .              1320
-----
1321 gcgaccgcctctctgtccccgccccctgtcattctacctccaacggcgtgcaggctccc 1380
    .              1340             .              1360             .              1380

```

### DNA Parallel

1381 gcgcccgcggcgctccgcggccccaccagacacacccctgcgacctgccccctcatca 1440  
                  .                       .                       .  
                    1400                    1420                    1440

1441 cggtcacgcacatggctgcgcccgcgcggcgccccccgaccccccgatccccggaacgtcc 1500  
 . 1460 . 1480 . 1500

1501 gccgccatccaccgccccgcagcccccgaggcccgaggaccccgacccccgcgcgcgcgc 1560  
 . 1520 . 1540 . 1560

1561 cgccgtagcgccgctctgtgcgggggggctggggggcgcgggggaaggccacgccccct 1620  
 . 1580 . 1600 . 1620

1621 ccactttttccaggaatgcgcggccccccatgccccacccccgatgccccccgggggtc 1680  
 . 1640 . 1660 . 1680

1681 tttcctgggggggcctttgtcagggttgctgtgcacccacgggcccccttggtcacgctt 1740  
 . 1700 . 1720 . 1740

1741 ggctgggtcaccccgaggcccggtcacgtttgccggggtcagccctggggcccttcgt 1800  
 . 1760 . 1780 . 1800

1801 cacgctccattgcccgtccccgcccccttgggggcttggtggcgcgccggggggcgc 1860  
 . 1820 . 1840 . 1860

1861 gaccgcctctctgtccccgccccctgtcattctaccctccaacggcgtgcaggctccgc 1920  
 . 1880 . 1900 . 1920

1921 gccgccccccgctccccgccccaccagacacacccccctgcgaccctgccccctcatcacg 1980  
 . 1940 . 1960 . 1980

1981 gtcacgcatggctgcgcccgcgcgcccccgacccccgatccccggaacgtccgc 2040  
 . 2000 . 2020 . 2040

2041 cgccatccaccgcccgcagccccggagcccgcggaaccccgacccccgccgcccgcgcg 2100  
                  .                    .                    .

2101 ccgtagcgcgcgtctgtgcggggggctggggggccgcgggggaaggccacgccccctcc 2160  
 . 2120 . 2140 . 2160





2941 gttatgagtgactggactggaggagccctccttgctcctctattcctttgctctcatgctt 3000  
 . 2960 . 2980 . 3000

3001 ataattataatTTTgatcatctttatctttcagaagagacTTTctctgtccacttggagcc 3060  
 . 3020 . 3040 . 3060

3061 ctttgatatactcctactgatgagtaagtattacacctttgccccacacccccctttccct 3120  
 . 3080 . 3100 . 3120

3121 tactcttctcttctctaacgcactttctctctttcccccagtcaccctcctgctcatcgct 3180  
 . 3140 . 3160 . 3180

3181 ctctggaatttgcacggacaggcattgttccttgggaattgtgctgttcattcttcgggtgc 3240  
 . 3200 . 3220 . 3240

3241 ttacttggtgaagatctaacattccctaggaattattttaccacacccccacttttccaacc 3300  
 . 3260 . 3280 . 3300

3301 ctaacactcttttttcaacgcagtcttaggtatctggatctacttattggagatgctctg 3360  
 . 3320 . 3340 . 3360

3361 gcgacttgggtgccaccatctggcagcttttggccttcttcttagccttcttcttagacct 3420  
 . 3380 . 3400 . 3420

3421 catcctgctcattattgctctctatctacaacaaaactggtggactctattggttgatct 3480  
 . 3440 . 3460 . 3480

3481 cctttggctcctcctgtttctggcgattttaatctggatgtattaccatggacaacgaca 3540  
 . 3500 . 3520 . 3540

3541 cagtgatgaacaccaccacgatgactccctcccgccaccctcaacaagctaccgatgattc 3600  
 . 3560 . 3580 . 3600

3601 tggccatgaatctgactctaactccaacgagggcagacaccacctgctcgtgagtggagc 3660  
 . 3620 . 3640 . 3660

3661 cggcgacggaacccccactctgctctcaaaacctaggcgacactggaggtggctctgacaa 3720  
 . 3680 . 3700 . 3720

3721 tggccacaggaccctgacaacactgatgacaatggcccacaggaccctgacaacactga 3780  
 . 3740 . 3760 . 3780

3781 tgacaatggcccatgacccgctgcctcaggaccctgacaacactgatgacaatggccc 3840  
 . 3800 . 3820 . 3840

3841 acaggaccctgacaacactgatgacaatggcccacatgaccgctgcctcatagccctag 3900  
 . 3860 . 3880 . 3900

1 -----catgatggcagcgcggt----- 18  
 x||||| | |  
 3901 cgactctgctggaaatgatggaggccctccacaattgacggaagagggttgaaaacaaagg 3960  
 . 3920 . 3940 . 3960

3961 aggtgaccaggggcccgccctttgatgacagacggaggcgggcggtcatagtcatgattccgg 4020  
 . 3980 . 4000 . 4020

4021 ccatggcggcggtgatccacaccttctacgctgcttttgggttcttctggttcgggtgg 4080  
 . 4040 . 4060 . 4080

4081 agatgatgacgacccccacggcccagttcagctaagctactatgactaaccttttctttac 4140  
 . 4100 . 4120 . 4140

4141 ttctaggcattaccatgtcataggcttgccctgactgactctccctccattttactgggaat 4200  
 . 4160 . 4180 . 4200

4201 gccttagctaatcaccttaactggcacacactcccttagccacactgtctgttaggctg 4260  
 . 4220 . 4240 . 4260

4261 aaaagccacattcatatttctatttcaaaacaaggggaaaggaggacatgcgagaattggc 4320  
 . 4280 . 4300 . 4320

4321 agacacctttaccagcccttaaacacaccacacaggtagcaaggaccgcggcggttgccag 4380  
 . 4340 . 4360 . 4380

4381 actccgccaccaacgcccctgcgttgaacccacccctcctacacacatcagacctctgca 4440  
 . 4400 . 4420 . 4440

4441 caacacaactaccaggcagatgaggcccttacttccacaggggtactggcataccagcgg 4500  
 . 4460 . 4480 . 4500

```
-----  
4501 gggaccacatacatccctgtctcccacccagtaactccagcaactttgctttccatcttg 4560  
      .               4520               4540               4560  
      19 -----gccaagcc----- 27  
      | | | | | | | |  
4561 tgccaatacacatttggattcagcccaagccacacctaactcatgccagcagaggcagga 4620  
      .               4580               4600               4620  
-----  
4621 acacctgttggtgacacattctttgcgcataagcactttaatccctctctcacaccaga 4680  
      .               4640               4660               4680  
-----  
4681 aactaagagctagcccaaacctccacacctgtcctcgctcatctttccacattcctctg 4740  
      .               4700               4720               4740  
-----  
4741 gccttctttccttgtccttactgtataaaagtccacgaaaacagctgtgcctcactctcg 4800  
      .               4760               4780               4800  
-----  
4801 agatggtacacgtcctggagcgtgctttgctagagcagcagtcctctgcctgcggcctgc 4860  
      .               4820               4840               4860  
-----  
4861 ccggtcttctctacggagaccaggcctagccaccctgccccgaggaccagacgtcagca 4920  
      .               4880               4900               4920  
-----  
4921 gactaagactactcctgggtggtactctgtgtcctgtttggacttttatgcctgctcctca 4980  
      .               4940               4960               4980  
-----  
4981 tctaagaagccaccatgcgaccgggtagaccactggctggattctacgtactctccgcc 5040  
      .               5000               5020               5040  
-----  
5041 gttccttcagaagaatgtccaaaagggtcaaagaacaaggccaagaaggagcgtgtccccg 5100  
      .               5060               5080               5100  
-----  
5101 tggaggaccgcccaccgactccgatgccaccagccagcagactgatccgcagaaacgcgt 5160  
      .               5120               5140               5160  
-----  
5161 tgggaggaggcgtccgccccgatgcggaggactgcatccaacgcttccacccccctggagc 5220  
      .               5180               5200               5220  
-----  
5221 cagcgtggtgggtgtcaacaaagaactttgacctgttgctcctgagatgtgaattgggat 5280  
      .               5240               5260               5280
```

5281 ggtgtggataaacatctcccgctagatggcgcccttattattgatgtgacttgtgatgcaa 5340  
 . 5300 . 5320 . 5340

5341 taaataaaagtacagatagatggcactcttaccttctctgcccgttcttcgtatatgt 5400  
 . 5360 . 5380 . 5400

5401 gttgagatgagtcacccgtggagagtagggagggggagggagcccgtcattcccgctcgt 5460  
 . 5420 . 5440 . 5460

5461 gttgcaatcccaagtacagactttgatcttgggttcttagtggttgatagtcaggatgac 5520  
 . 5480 . 5500 . 5520

5521 ggtcgccattgccccaatatgggtcctcataaggcggtgggggctcttcattagattcac 5580  
 . 5540 . 5560 . 5580

5581 gttcctcatcgttcggtgggggtgggggtgttcccagaagagccagaagcagatggatatt 5640  
 . 5600 . 5620 . 5640

5641 gggagttgtttccgccatcgtaacctatccggatccccgccggggctagggggacccgcgc 5700  
 . 5660 . 5680 . 5700

5701 ccattggcaccatttctagggacccccatagctgcagcagcgactgcaaaaccggcctgaa 5760  
 . 5720 . 5740 . 5760

5761 atgctctctgagaaacaaggcgagagggatttacgccccagcaagcttcaggcaaaaataa 5820  
 . 5780 . 5800 . 5820

5821 cctgtgatgctgaaaccgcggcgctacaggatcaccccaattggcaagacctgggcggag 5880  
 . 5840 . 5860 . 5880

5881 ttctgagagcccaggggtctcgtgcagggtgtccccggggaattgttctccctgatcacc 5940  
 . 5900 . 5920 . 5940

28 ---cacc----- 32

5941 gcccacccccgttttctccaaaccagcagctgtgacaatcttcacacactgctgctgtca 6000  
 . 5960 . 5980 . 6000

6001 cctggaactattttcccacgggtgcccttccgcccattttcccacgagtcgcgaggctatc 6060  
 . 6020 . 6040 . 6060

-----  
 6061 caccgcgaatgccaccccccaatgccacactaaaacaagggtgaaataggcaagtgcgtt 6120  
                                 6080                                6100                                6120  
 -----

6121 tattgcgacaagtatccagaaacataaacccccgtgggcttctctcttgcatttttccca 6180  
                                 6140                                6160                                6180  
 -----

6181 acgcaggtcactggcaggtgccagggcttgggaagtgcaggtcaacagcaacagagagg 6240  
                                 6200                                6220                                6240  
 -----

6241 ctcccatccttttcttccataacaccgccatttgccgcagttggtgcgggctccacgccc 6300  
                                 6260                                6280                                6300  
 -----

6301 tcgggcatgagccactggacgtggggatggggaaatgcattcacggtgcatgtcacagta 6360  
                                 6320                                6340                                6360  
 -----

6361 aggacagagaagtctggaactgagaccttctggagtggaacagacagcgtagaggcttc 6420  
                                 6380                                6400                                6420  
 -----

6421 accacgtcaggtgttctctgcttggtagacctcggtctcgcccagtttcatgcggcacagg 6480  
                                 6440                                6460                                6480  
 -----

6481 tagttgccgtcatgggagatggtggcagcggtgactactaaaaagaagggtgttggcactt 6540  
                                 6500                                6520                                6540  
 -----

6541 ctgtggatatcaaagaagcccctgaaaggccactctataaagatgacatcggtgcatg 6600  
                                 6560                                6580                                6600  
 -----

6601 cgcccaataagcacctgctctctctctgggcccagtttaaacagctgacctcaatctct 6660  
                                 6620                                6640                                6660  
 -----

33 -----tccag----- 37  
                                 |||||  
 6661 ggaccgaggctcacctctctccagtaggaggtcagggtgactcgctcacccaagaaagcg 6720  
                                 6680                                6700                                6720  
 -----

6721 gtgacagcctggccggcgccacacaggaggccaacaggaggagctgagcgatgaacctg 6780  
                                 6740                                6760                                6780  
 -----

6781 gccattgctctggactctctctcaccagggcctcgcttattatactattctgccacgcccc 6840  
                                 6800                                6820                                6840  
 -----

[illegible]

-----

---

-----

-----

-----

\_\_\_\_\_

40 .  
gggaggctggaggcggattt  
| | | | | | | | xxxxxxxxxxxx

\_\_\_\_\_

\_\_\_\_\_

---

-----

7561 agccagccgggtccgtgcggtacatctacttttaccacctggccacctacctcttagg 7620  
 . . 7580 . 7600 . 7620

[illegible]

.....

\_\_\_\_\_

[illegible]

100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

\_\_\_\_\_

-----

.....

-----

[illegible]

-----

.....

8341 agcgcggcatgcgagctggcggtagctacgcacccgccatgctggaggcctcgcacgag 8400  
 . 8360 . 8380 . 8400

```

58 -----tccaga- 63
      |||||
8401 accccggccagcctcaactacgactcgtggccctgtttgccgactgtgaggtccagag 8460
      .           8420           .           8440           .           8460
-----
8461 gcccggtgtggctgcgttacaccgatataatgccagcctggccccccacgtgtccacgcag 8520
      .           8480           .           8500           .           8520
-----
8521 atctttgccaccaattccgtcctctacgtctcgggggtctcgaagtcaaccggtcagggc 8580
      .           8540           .           8560           .           8580
-----
64 -----cagt----- 67
      |||||
8581 aaggagagtctctttaacagtttctacatgaccacggcctggggaccctgcaggagggg 8640
      .           8600           .           8620           .           8640
-----
68 -----ccctgcttc----- 77
      |||||
8641 acctgggaccctgcccggaccctgcttctcgggctggggggccagacgtgaccgga 8700
      .           8660           .           8680           .           8700
-----
8701 accaacggtccgggaaactacgctgtggagcacctgggtctatgcggcctccttctcgccc 8760
      .           8720           .           8740           .           8760
-----
8761 aaccttcttgcccgtatgcctactacctgcagttttgccaggacagaagagctctctg 8820
      .           8780           .           8800           .           8820
-----
8821 accccggtgccggagacgggcagctacgtggcggggggcggccgagtcctcatgtgctcg 8880
      .           8840           .           8860           .           8880
-----
8881 ctctgcgagggccggggcccggtgtgcctgaacacgctcttcttttaggctgagggac 8940
      .           8900           .           8920           .           8940
-----
8941 cgcttcccccccgctcatgtccacgcagcggaggggaccctatgtgatctcgggggcctcg 9000
      .           8960           .           8980           .           9000
-----
9001 ggctcctacaacgagacggactttttgggcaactttctcaacttcacgataaggaggac 9060
      .           9020           .           9040           .           9060
      80           100
78 -----cta-aatttcaag-agctgaaccagaa----- 102
      |||||
9061 gacgggcagcggccggacgacgagccccgctacacctactggcagctgaaccagaacctg 9120
      .           9080           .           9100           .           9120
-----
9121 ctggagcgggctgtctcgggctgggcatagacgctgaaggaaagctagagaaggagcccat 9180
      .           9140           .           9160           .           9180
```



9181 gggccgcgctgactttgtcaagatgttcaaggacgtggatgcgggcggtggacgccgaagtg 9240  
 . 9200 . 9220 . 9240

9241 gtccagtttatgaacagcatggccaagaacaacatcacctacaaggacctgggtcaagagc 9300  
 . 9260 . 9280 . 9300

9301 tgctaccacgtgatgcagtactcgtgcaacccctttgcgcagcccgccctgccccatcttc 9360  
 . 9320 . 9340 . 9360

103 -----taatctcccaatgatgtt 122  
 9361 acccagctgttttaccgctcactgctgaccatcctgcaggacatctccctgccatctgt 9420  
 . 9380 . 9400 . 9420

123   ttcgg-----127

9421 atgtgctatgagaatgacaaccccgaggcttggccagagccccccagagtggctaagggt 9480

                9440                    9460                    9480

9481 cactaccagacgctgtgcaccaacttttaggagcctggccatcgacaagggggtcctcacg 9540  
 . 9500 . 9520 . 9540

9541 gccaaggaggccaaggtggtgcatggggagcccacctgcgacctgccagacctggacgcg 9600  
 . 9560 . 9580 . 9600

9601 gccctgcagggccgggtgtacggccggcggtgcctgtgcgcatgtccaaggtgctgatg 9660  
 . 9620 . 9640 . 9660

9661 ctgtgccccaggaacatcaagatcaagaacagggtggtcttcacgggggagaatgccgcc 9720  
 . 9680 . 9700 . 9720

9721 ctccagaacagcttcatcaagtcactaccaggagggagaactacatcatcaacggggccc 9780  
 . 9740 . 9760 . 9780

9781 tacatgaaattcctcaacacctaccacaagaccctattcccggacactaagctctcaagc 9840  
 . 9800 . 9820 . 9840

9841 ctgtacctgtggcacaacttttccaggcgggcgctcgggtccctgtccccagcggggccagc 9900  
 . 9860 . 9880 . 9900

9901 gcggaggagtactctgacctggccctctttgtggacgggggctcccgggccacgaagag 9960  
 . 9920 . 9940 . 9960

[illegible]

```
-----
10741 tgggggtgggcccggggcagcccaggacaactttatcagcgtggccgagccgggtcagcacc 10800
      .           10760           .           10780           .           10800
      .
205 -----gcctcggc----- 212
      |||||
10801 gcgtcgcaggcctcggccgggctgctgctgggtggaggagggcagggtccgggggcaga 10860
      .           10820           .           10840           .           10860
      .
-----
10861 cgcaagcgccgtctggccaccgttctccccggactcgagggtctagagaccctggggcg 10920
      .           10880           .           10900           .           10920
      .
-----
10921 cgatgtcggggctgctggcgggcggtacagccaggtgtacgccctggcggttgagctga 10980
      .           10940           .           10960           .           10980
      .
-----
10981 gcgtgtgcacccggctggaccccgagctctggacgtggctgcgggtggtgcgcaacgccg 11040
      .           11000           .           11020           .           11040
      .
-----
11041 gcctgctggccgagctggaggccatcctccttccccgtttgagacggcagaatgaccgtg 11100
      .           11060           .           11080           .           11100
      .
-----
11101 catgcagcgccctgtccctggagctggtgcacctgctagagaactcgagagaggcctctg 11160
      .           11120           .           11140           .           11160
      .
-----
11161 ccgcgctgctcgcccctggtagaaagggtacccgggtcccgccctctccgtacccctcag 11220
      .           11180           .           11200           .           11220
      .
-----
11221 tcgcgctactctgtggagttttacggggggcataaagtcgatgtaagtttgtgcctaataa 11280
      .           11240           .           11260           .           11280
      .
-----
11281 atgacatagagattttaatgaagagaatcaatagcgtgttttattgcatgtctcacacca 11340
      .           11300           .           11320           .           11340
      .
-----
11341 tggggctggagagcctggaacgggcccctggatctgctgggcccgtttcggggcgtaagtc 11400
      .           11360           .           11380           .           11400
      .
-----
11401 ccatcccagaccgcgcctctacatcacctctgtgccctgctggcgctgtgtgggcgagc 11460
      .           11420           .           11440           .           11460
      .
-----
11461 tgatggttctgcccaccacggcaacccttccacggcagaggggaccacgtctcctgta 11520
      .           11480           .           11500           .           11520
```

-----  
 11521 accacctggcgggtgccgggtgaatccggagccgggtctcgggactgtttgagaatgaagtcc 11580  
                   11540                                  11560                                  11580  
 -----

11581 gccaggcgggggctcggggcacctgttggaggctgaggagaaggcgaggccggggcgccag 11640  
                   11600                                  11620                                  11640  
                   220                                  240

213 -----gccaacgcgc-catagacaagaggcagagagccagtgtggctggggct 259  
                   ||      ||      |      |      ||      ||      ||      ||

11641 aggagggcgcggtcccggggcccgggcgggcgccggaggcagag===== 11681  
                   11660                                  11680

260                  280                                  300  
 260 ggtgctcatgcacaccttggcgggtcatccgccaccccgctccagcaggctcaggccgcc 319  
 =====

320                  340                                  360  
 320 gcatccgctgggaccggggccttggcatcatcagcgccgctccacggccgtagcccagtcc 379  
 =====

380                  400                                  420  
 380 gcgacccccctctgtttcttcatctattagcagcctccgggcccgcgacttcgggggcgac- 438  
   |||||  
 11682 =====ggggcgacc 11690

-----  
 11691 agagcgctggacacctacaacgtcttctcgacagtcccccgagggtggcgagctctcg 11750  
                   11700                                  11720                                  11740  
 -----

11751 gagtcctctattggaactctggcgcccatgctatcggtgcaacggggcagggggagggt 11810  
                   11760                                  11780                                  11800  
 -----

11811 ggcggccattcccgcctctctgcccgtgtttgcccgaggcgctgcctggccctggtgcgg 11870  
                   11820                                  11840                                  11860  
 -----

11871 ggggcctgcgaggaggcgctggcgggggcaaggctgactcacctgtttgacgccgtggct 11930  
                   11880                                  11900                                  11920  
 -----

11931 cccggggccacggagcggtcttctgcggcggggtctacagctcctcggggcagcggtg 11990  
                   11940                                  11960                                  11980  
                   440                                  460

439 -----tgccgcgcctccgcgcgcagccgtcgataaccggg--- 475  
                   |||||  
 11991 gaggcgctgaaggcggaactgcgcgcgccttcacggc=gcaccccag=taccggggcca 12048  
                   12000                                  12020                                  12040  
 -----

12049 tcctgcaaaagaggaacgagctgtacacgcgggtcaaccgagccatgcagcggttggggc 12108  
                   12060                                  12080                                  12100

```

-----
12109 gaggcgaggaggagggcgctcccgggagagcccggaggtgccccggccggctggggcacgag 12168
      .      12120      .      12140      .      12160

476 -----tc 477
      |
12169 agccccggcccgctccggcgccctctcggacgcgctcaagcgcaaggagcagtacctgcgcc 12228
      .      12180      .      12200      .      12220
480
478 aggtggc----- 484
      |||||
12229 aggtggccaccgaggggtctggccaagctgcagtcctgcctggcgcaacagagcgagaccc 12288
      .      12240      .      12260      .      12280
      .
485 -----gggggacaacccc-ac-gacacc-g-ccc-ca 511
      .      .      .      .      .      .      .      .      .      .
12289 tgaccgagaccctgtgcctgcgcgtctggggggacgtggtctactgggagctggcccgca 12348
      .      12300      .      12320      .      12340
      520
512 cgcggggc-acgtaagaaacagtagccc----- 538
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
12349 tgcgcaaccacttcctctacagacgggccttcgtctcgggtccctgggaggacaggcgcg 12408
      .      12360      .      12380      .      12400

-----

12409 ccggcgaggggtgccgcctttgagaattc 12436
      .      12420

```

% Identity = 1.9 (243/12628)

///

9310-13DVCTDV SEQ ID NO 3.xdna x Bankier et al. EcoRI Dhet fragment.xdna => DNA Parallel

DNA sequence 1038 bp atgctatcaggt ... cgcgtggcttga linear

DNA sequence 12436 bp gaattctcaaag ... tgtttagaattc linear

Method: Blocks (Martinez)  
Layout: Standard  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Translation: Off

Alignment 9. Comparison of nucleotide sequence  
of SEQ ID NO:3 with the nucleotide sequence of  
Fig. 2 of Bankier et al.

```

-----
1  gaattctcaaaggcggcaccctcgccggcgcgccctgtcctcccagggaacccgagacgaag 60
    .                20                .                40                .                60
-----
61  gcccgtctgtagaggaagtgggttgcgcattgcggggccagctcccagtagaccacgtccccc 120
    .                80                .                100               .                120
-----
121 cagacgcgcaggcacagggtctcggtcagggtctcgtctgttgcgccaggcaggactgc 180
    .                140               .                160               .                180
-----
181 agcttgggccagaccctcggtggccacctggcgccaggtactgctccttgcgcttgagcgcg 240
    .                200               .                220               .                240
-----
241 tccgagagggcgccgggacggggccgggctctcgtgccccagccggccggggcacctccggg 300
    .                260               .                280               .                300
-----
301 ctctcccgggacgcctcctcctcgccctcgcccaaccgctgcatggctcggttgagccgc 360
    .                320               .                340               .                360
-----
361 gtgtacagctcgttcctcttttgcaggatggccccgggtactgggggtgcgccgtgaaggcg 420
    .                380               .                400               .                420
-----
421 gcggcgagctccgccttcagcgccctccaccgcgtcgcccgaggagctgtagaccccgccg 480
    .                440               .                460               .                480
-----
481 cagaagagccgctccgctggccccgggagccacggcgctcaaacagggtgagtcagccttgcc 540
    .                500               .                520               .                540
-----
541 cccgccagcgccctcctcgccaggcccccccgaccaggggccaggcgacgctcccgggcaaac 600
    .                560               .                580               .                600

```

```

601 agggcagagagggcggggaatggccgccaccctccccctgccccgttgccacgatagcatgg 660
    .               620               .               640               .               660
-----

661 ccgccagagttccaatagaggagctccgagagctccgccacctccgggggcactgtcgag 720
    .               680               .               700               .               720
-----

721 aagacgttgtaggtgtccagcgctctgggtcgccccctctgcctccggccgccccggggccc 780
    .               740               .               760               .               780
-----

781 gggaccgcgcctcctcttgggcgcccgccggcctcgcccttctcctcagcctccaacaggtgc 840
    .               800               .               820               .               840
-----

841 ccgagccccgcctggcggaacttcattctcaaacagtcgccgagaccggctccggattcacc 900
    .               860               .               880               .               900
-----

901 ggcaccgccaggtgggttacaggagacgtgggtccccctctgccttggaaggggttgccgtgg 960
    .               920               .               940               .               960
-----

961 ttgggcagaaccatcagctcgccacacagcgccagggcacagaggtgatgtagagg 1020
    .               980               .               1000               .               1020
-----

1021 cgcgggtctgggatgggacttacgccccgaaagcgggccagcagatccagggcccgttcc 1080
    .               1040               .               1060               .               1080
-----

1081 aggctctccagccccatgggtgtgagacatgcaataaaacacgctattgattctcttcatt 1140
    .               1100               .               1120               .               1140
-----

1141 aaaatctctatgtcattttattaggcacaaaacttacatcgactttatgcccccgtaaaac 1200
    .               1160               .               1180               .               1200
-----

1201 tccacagagtacgcgactgaggggggtacggagaggcgggaccgggtaccctttctacca 1260
    .               1220               .               1240               .               1260
-----

1261 ggggcgagcagcgcggcagaggcctctctcgagttctctagcaggtgcaccagctccagg 1320
    .               1280               .               1300               .               1320
-----

1321 gacagggcgctgcatgcacggtcattctgcctctcaaacgggggaaggaggatggcctcc 1380
    .               1340               .               1360               .               1380

```

\_\_\_\_\_

---

\_\_\_\_\_

---

---

---

=====

.....

---

\_\_\_\_\_

-----

---

2101 ccataaaagtatcccaggttcccggcctggaatatctggttggtgcggttgacccccgtgt 2160  
 . 2120 . 2140 . 2160



**Table 1**

=====

1 atggtatgag 10

\_\_\_\_\_

\_\_\_\_\_

**100**

.....

.....

.....

-----

[illegible]

-----

2881 ccttggcctccttggccgtgaggaccccttgtcgatggccaggctcctaaagttggtgc 2940  
 . 2900 . 2920 . 2940

.....

---

-----

-----

-----

-----

---

-----

---

-----

---

-----

3661 agcgggcaagaaggttgggcgagaaggaggccgcatagaccagggtgctccacagcgtagt 3720  
 . 3680 . 3700 . 3720

-----  
 3721 ttcccggaacggttggttccgggtcacgtctggtccacccagcccgagaagcaggggtcggc 3780  
                                   3740                                  3760                                  3780  
 -----

3781 ggcaggggtcccaggtcccctcctgcaggggtccccaggccgtgggtcatgtagaaactgt 3840  
                                   3800                                  3820                                  3840  
 -----

3841 taaagagactctccttgccctgaccggttgacttcgagacccccgagacgtagaggacgg 3900  
                                   3860                                  3880                                  3900  
 -----

11 -----gta 13  
 3901 aattggtggcaaagatctgcgtggacacgtggggggccaggctggcattatatcgggtgta 3960  
                                   3920                                  3940                                  3960  
                                   |||||

14 acgc----- 17  
 3961 acgcagccacacggggcctctggaccctcacagtcggcacaacagggggccacgagtcgtagt 4020  
                                   3980                                  4000                                  4020  
 -----

4021 tgaggctggccggggtctcgtgcgaggcctccagcatggcgggtgcgtagctcaccgcca 4080  
                                   4040                                  4060                                  4080  
 -----

4081 gctcgcattgcgcgctgtccacaatcattaaggctcccaggtccgggtgactgatgggtg 4140  
                                   4100                                  4120                                  4140  
 -----

4141 aggctgggaactccttgagggggggccaccttgggccaccttggcctggtcctgcaggctct 4200  
                                   4160                                  4180                                  4200  
 -----

4201 gcttctccagcagctccaccagcttgcccacccgtcggacgcgcagcgcctgcgccagcc 4260  
                                   4220                                  4240                                  4260  
 -----

4261 cgggtgtacagcgcctcgtgcattgcagcggctgaggtccgagttgtaaaactggcggagct 4320  
                                   4280                                  4300                                  4320  
 -----

4321 ggggcacgcctcttggaacacctccttgctgtagagcgggaccctaacgctcgcagact 4380  
                                   4340                                  4360                                  4380  
 -----

4381 gccccaccgctacctcctgttttaacgatggaatggccaccagggtttccgctgtagagtc 4440  
                                   4400                                  4420                                  4440  
 -----

4441 gctccttgaaggcctcgggttattgccaccgccccccaggtaggcagagggatctagccctt 4500  
                                   4460                                  4480                                  4500  
 -----

5221 gccgcgaccagtagtcgtagtggtcgttgtagacactgcgcgcaggacgctgatgatgagcc 5280  
 . 5240 . 5260 . 5280

[illegible][illegible]

\_\_\_\_\_

100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

.....

[illegible]

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

1999 1998 1997 1996 1995 1994 1993 1992 1991 1990 1989 1988 1987 1986 1985 1984 1983 1982 1981 1980 1979 1978 1977 1976 1975 1974 1973 1972 1971 1970 1969 1968 1967 1966 1965 1964 1963 1962 1961 1960 1959 1958 1957 1956 1955 1954 1953 1952 1951 1950 1949 1948 1947 1946 1945 1944 1943 1942 1941 1940 1939 1938 1937 1936 1935 1934 1933 1932 1931 1930 1929 1928 1927 1926 1925 1924 1923 1922 1921 1920 1919 1918 1917 1916 1915 1914 1913 1912 1911 1910 1909 1908 1907 1906 1905 1904 1903 1902 1901 1900

.....

6001 aacacctgagcgtggtgaagcctctaacgctgtctgtccactccgaaagggtctcagttcc 6060  
 . 6020 . 6040 . 6060

-----  
 6061 cagacttctctgtccttactgtgacatgcacogtgaatgcatttccccatccccacgtcc 6120  
                                 6080                                6100                                6120  
 -----

6121 agtggctcatgcccgagggcggtggagcccgcaccaactgcggcaaatggcggtgttatga 6180  
                                 6140                                6160                                6180  
 -----

6181 aggaaaaggatgggagccctctctgttgctgttgacctgtcacttcccaagccctggcacc 6240  
                                 6200                                6220                                6240  
 -----

6241 tgccagtgcacctgcgttgggaaaaatgacaaggaggaagcccacggggtttatgtttctg 6300  
                                 6260                                6280                                6300  
 -----

6301 gatacttgtcgcaataaacgcacttgccattttcaccttgttttagtgtggcattggggg 6360  
                                 6320                                6340                                6360  
 -----

6361 ggtggcattgcgggtggatagcctcgcgactcgtgggaaaatgggcggaagggcacctg 6420  
                                 6380                                6400                                6420  
 -----

6421 ggaaaatagttccagggtgacagcagcagtgtgtgaagattgtcacagctgctggtttgga 6480  
                                 6440                                6460                                6480  
 -----

6481 gaaaacgggggtgggcggtgatcagggagaacaattccccggggacacctgcacgagacc 6540  
                                 6500                                6520                                6540  
 -----

6541 cctgggctctcaggaactccgcccagggtcttgccaattgggggtgatcctgtagcgcgcg 6600  
                                 6560                                6580                                6600  
 -----

6601 gtttcagcatcacagggttattttgcctgaagcttgctggggcgtaaatccctctgcctt 6660  
                                 6620                                6640                                6660  
 -----

6661 gtttctcagagagcatttcaggccggttttgcagtcgctgctgcagctatggggtcccta 6720  
                                 6680                                6700                                6720  
 -----

6721 gaaatggtgccaatgggcgcgggtccccctagccccggcggggatccggatgggtacgat 6780  
                                 6740                                6760                                6780  
 -----

6781 ggcggaaacaactcccaatatccatctgcttcttggtcttctgggaacacccccaccca 6840  
                                 6800                                6820                                6840  
 -----

7561 tccgtagaagagccgggcaggccgcaggcagaggactgctgctctagcaaagcacgctcc 7620  
 . 7580 . 7600 . 7620

7621 aggacgtgtaccatctcgagagtgaggcacagctgttttcgtggacttttatacagtaag 7680  
 . 7640 . 7660 . 7680

7681 gacaaggaaagaaggccagaggaatgtggaaagatgagcgaggacaggtgtggaggtttt 7740  
 . 7700 . 7720 . 7740

7741 gggctagctcttagtcttctgggtgtgagagagggattaaagtgcttatgcgcaaagaatg 7800  
 . 7760 . 7780 . 7800

7801 tgtcaacaacaggtgttcctcgccctgtgctggcatgagttaggtgtggcctgggctgaatc 7860  
 . . 7820 . 7840 . 7860

7861 caaatgtgtattggcacaagatggaaagcaaagtgtgctggagttactgggtgggagacag 7920  
 . 7880 . 7900 . 7920

7921 ggatgtatgtgggtcccccgctgggatgccagtaacctgtggaagtaagggggcctcatctg 7980  
 . 7940 . 7960 . 7980

7981 cctggtagttgtgttggtgcagaggtctgatgtgtgtaggaggggtgggttcaacgcaggg 8040  
 . 8000 . 8020 . 8040

8041 gcgttggtggcgagtcctggcaacgcccgggtccttgctacctgtgtggtgtgttaaggg 8100  
 . 8060 . 8080 . 8100

8101 ctgggtaaagggtgtctgccaattctcgcatgtcctcctttcccttgttttgaaatagaa 8160  
 . 8120 . 8140 . 8160

8161 tatgaatgtggccttttcagcctagacagacagtgtggctaaggaggagtgtgtgccagttaa 8220  
 . 8180 . 8200 . 8220

8221 ggtgattagctaaggcattcccagtaaatggagggagagtcagtcaggccaagcctatgac 8280  
 . 8240 . 8260 . 8280

8281 atggtaatgcctagaagtaaaagaaaggttagtcatagtagcttagctgaactgggccgtg 8340  
 . 8300 . 8320 . 8340

8341 ggggtcgtcatcatctccaccggaaccagaagaacccaaaaagcagcgtaggaagggtgtgg 8400  
 . 8360 . 8380 . 8400



8401 atcaccgcccgcgatggccggaatcatgactatgaccgcccgcctccgtctgtcatcaaagg 8460  
 . 8420 . 8440 . 8460

8461 cggggccctgggtcacctcctttgttttcaacctcttcgcgtcaattgtggaggggcctccatc 8520  
 . 8480 . 8500 . 8520

8521 atttccagcagagtcgctagggcctatgaggcagcgggtcatgtggggccattgtcatcagt 8580  
 . 8540 . 8560 . 8580

8581 gttgtcagggtcctgtgtgggccattgtcatcagtggttgtcagggtcctgaggcagcgggtc 8640  
 . 8600 . 8620 . 8640

8641 atgtggggccattgtcatcagtgttgtcagggtcctgtggggccattgtcatcagtgttgtc 8700  
 . 8660 . 8680 . 8700

8701 agggctcctgtggggccattgtcaggaccacctccaggtgcgcctaggttttgagagcagag 8760  
 . 8720 . 8740 . 8760

8761 tgggggtccgctcgccggctccactcacgagcaggtggtgtctgccctcgttggagttaga 8820  
 . 8780 . 8800 . 8820

8821 gtcagattcatggccagaatcatcggtagcttgttgagggtgcgggagggagtcacgtg 8880  
 . . . 8840 . 8860 . 8880

8881 gtggtgttcatactgtgtcgttgtccatggtaatacatccagattaaaatcgccagaaa 8940  
 . 8900 . 8920 . 8940

8941 caggaggagccaaaggagatcaaccaatagagtccaccagttttgttgtagatagagagc 9000  
 . 8960 . 8980 . 9000

9001 aataatgagcaggatgaggtctaggaagaaggctaggaagaaggccaaaagctgccagat 9060  
 . 9020 . 9040 . 9060

9061 ggtggcaccaagtcgccagagcatctccaataagtagatccagatacctaagactgcgtt 9120  
 . 9080 . 9100 . 9120

9121 gaaaaaaagagtgttagggttgaaaaagtgggggtgtggtaaataattcctaggggaatgtt 9180  
9140 . 9160 . 9180

9743 gggat===== 9747

360 380 400

345 ccctccggcctacttttggttgccgggctcttttgccccccctccaccgtgcctccttac 404

=====

420

405 tacggattcccacttgccgg----- 424

9748 =====ttgccccgggtctgccccgggagcagtcaggggtacagatttcccgaaagcg 9794

9760 9780

-----

9795 gcggtgtgtgtgtgcatgtaagcgtagaaaggggaagtagaaagcgtgtgtttgtgttag 9854

9800 9820 9840

440 460 480

425 cagactacgtccccg-ctccctcgcgatccaacaagcgg-aaaagagacccccgaggagga 482

9855 aaaagcgggtccccgggggcaagctgtggga=atgcggtggccaagtgcacacagga=aa 9912

9860 9880 9900

500 520 540

483 tgaagaagcgccccgggctat-tccccggggaggacgccaccctctaccgcaaggacatag 541

9913 tggaagcgagtcgccccgaatcagaagggggag===== 9944

9920 9940

560 580 600

542 cgggcctctccaagagtgtgaatgagttacagcacacgctacaggccctgcgccgggaga 601

=====

620 640 660

602 cgctgtcctacggccacaccggagtcggatactgccccagcagggccctgctacaccc 661

=====

680 700 720

662 actcggggccttacggatttcagcctcatcaaagctacgaagtgcccagatacgtccctc 721

=====

740 760 780

722 atccgccccaccaccaacttctcaccaggcagctcaggcgcagcctccacccccgggca 781

9945 =====tgcgtagtgtgtgtgggaagcggcagtgtaatctgcaca 9982

9960 9980

800 820

782 cacaggcccccgaaagcccactgtgtggccgagtcacgatccctgaggcgggag----- 835

9983 aagaggcgccccgcgcgaacgt=tgggag=gtcggtg=gcggcaggcgggaggccgtg 10038

10000 10020

-----

10039 ctttaggggggttcaggtgaggcaaggctgtggggtaaccgtaggggaggcgggtgaggc 10098

10040 10060 10080

840

836 -----cagccgggaa-ctc 848

10099 ggctaagagggttaagggtcgccgggtgacgaagcagcagacggcggatatgggaatttc 10158

10100 10120 10140

860 880 900

849 tggacccccg-ggaggacaccaaccctca-gcagccac-caccgag-ggc--ca-cca-c 900

10159 agaattgaggtggcggattcaggcgaaaagggtgtgggctgtgcgagtgatgagggcagg 10218

10160 10180 10200

```

          920          940          960
901  cgcggaaagaaactgggtgcaggcctctgcgtccggagtggtcagtcctaaggagcccacc 960
    ||||| |xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
10219 cgcggaaagtcgctgcggccttgctggggcatggggggccgcgcattcctggaaaaagtgg 10278
    10220          10240          10260
          980          1000          1020
961  accccaaggccaagtctgtgtcagcccacctaagtccatcttttgcgaggaattgctg 1020
    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
10279 agggggcgtggccttcccccgcgccccccagccccccgcacagagcggcgctacggcg 10338
    10280          10300          10320

1021  aataaacgcgtggcttga----- 1038
    xxxxxxxxxxxxxxxxxxxxxxx
10339 ggcggggcggcgggggggtcggggtccgcggggtccgggggctgcgggcggtggatggcggc 10398
    10340          10360          10380

-----

10399 ggacgttccgggggatcgggggggtcggggggcgccgcgcggggcgagccatgcgtgaccg 10458
    10400          10420          10440

-----

10459 tgatgagggggcagggtcgcagggggtgtgtctggtgggggcgggagcggggggcgggcg 10518
    10460          10480          10500

-----

10519 gggagcctgcacgccgttgagggtagaaatgacagggggcggggacagagaggcggtcgc 10578
    10520          10540          10560

-----

10579 gcccccgccgcgccagccaagcccccaagggggcggggagcgggcaatggagcgtgac 10638
    10580          10600          10620

-----

10639 gaaggggccccagggtgacccccggcaaactgacccggggctccggggtgacccagccaa 10698
    10640          10660          10680

-----

10699 gcgtgaccaagggggcccggtgggtgacacaggcaaccctgacaaaggccccccaggaaaga 10758
    10700          10720          10740

-----

10759 cccccggggggcatcgggggggtggggcatggggggccgcgcattcctggaaaaagtggag 10818
    10760          10780          10800

-----

10819 ggggcgtggccttcccccgcgccccccagccccccgcacagagcggcgctacggcggg 10878
    10820          10840          10860

-----

10879 cgggcggcggggggtcgggggtccgcggggtccgggggctgcgggcggtggatggcgggcg 10938
    10880          10900          10920

-----

10939 acgttccgggggatcgggggggtcggggggcgccgcgcggggcgagccatgcgtgaccgtg 10998
    10940          10960          10980

```

-----  
10999 atgagggggcaggggtcgcaggggggtgtgtctggtggggggcgggagcggggggcggcgcgg 11058  
11000 . 11020 . 11040 .

-----  
11059 gagcctgcacgccgttgagggttagaatgacagggggcggggacagagaggcggtcgcgc 11118  
11060 . 11080 . 11100 .

-----  
11119 ccccgccgcgccagccaagcccccaagggggggcggggagcgggcaatggagcgtgacga 11178  
11120 . 11140 . 11160 .

-----  
11179 agggccccaggggtgaccccggaacgtgacccggggctccgggggtgacccagccaagc 11238  
11180 . 11200 . 11220 .

-----  
11239 gtgaccaagggggcccggtgggtgacacaggcaaccctgacaaaggccccccaggaaagacc 11298  
11240 . 11260 . 11280 .

-----  
11299 cccgtggggcatggggggccgcgcattcctggaaaaagtggagggggcgtggccttcccc 11358  
11300 . 11320 . 11340 .

-----  
11359 cgcgccccccagccccccgcacagagcggcgctacggcgggcgggcgggcggggggtcg 11418  
11360 . 11380 . 11400 .

-----  
11419 ggggtccgcgggctccgggggctgcgggcgggtggatggcgggcgacgttccggggatcggg 11478  
11420 . 11440 . 11460 .

-----  
11479 ggggtcggggggcgccgcgcggggcgagccatgcgtgaccgtgatgagggggcagggtcg 11538  
11480 . 11500 . 11520 .

-----  
11539 caggggggtgtgtctggtggggggcgggagcggggggcggcgcgggagcctgcacgccgttg 11598  
11540 . 11560 . 11580 .

-----  
11599 gagggtagaatgacagggggcggggacagagaggcggtcgcgcccccgccgcgccagcc 11658  
11600 . 11620 . 11640 .

-----  
11659 aagcccccaagggggggcggggagcgggcaatggagcgtgacgaagggccccagggtgac 11718  
11660 . 11680 . 11700 .

-----  
11719 cccggcaaacgtgacccggggctccgggggtgacccagccaagcgtgaccaagggggcccg 11778  
11720 . 11740 . 11760 .

-----  
 11779 ggggtgacacaggcaaccctgacaaaggccccccaggaaagacccccggggggcatcgggg 11838  
 11780 . 11800 . 11820

-----  
 11839 ggtggggcatggggggccgcgcattcctggaaaaagtggagggggcggtggccttcccccg 11898  
 11840 . 11860 . 11880

-----  
 11899 cgccccccagccccccgcacagagcggcgctacggcgggcgggcgggggggtcggg 11958  
 11900 . 11920 . 11940

-----  
 11959 gtccgcggggtccgggggctgcgggcggtggatggcgggcgacgttcgggggatcggggg 12018  
 11960 . 11980 . 12000

-----  
 12019 ggtcggggggcgccgcgcgggcgcagccatgcgtgaccgtgatgagggggcagggtcgca 12078  
 12020 . 12040 . 12060

-----  
 12079 gggggtgtgtctggtggggcgggagcggggggcggcgcgggagcctgcacgccgttgga 12138  
 12080 . 12100 . 12120

-----  
 12139 gggtagaatgacagggggcggggacagagaggcggtcgcgcccccgccgcgccagccaa 12198  
 12140 . 12160 . 12180

-----  
 12199 gcccccaagggggcggggagcgggcaatggagcgtgacgaagggccccagggtgacct 12258  
 12200 . 12220 . 12240

-----  
 12259 cggcaaacgtgaccgggggtccggggtgaccagccaagcgtgaccaagggggcccggtgg 12318  
 12260 . 12280 . 12300

-----  
 12319 gtgacacaggcaaccctgacaaaggccccccaggaaagacccccggggggcatcgggggg 12378  
 12320 . 12340 . 12360

-----  
 12379 ggtgttgcgggggcatgggggggtcggtatttcgccttattgccctgtttagaattc 12436  
 12380 . 12400 . 12420

% Identity = 1.6 (210/12958)

///

9310-13DVCTDV SEQ ID NO 3.xdna x Bankier et al. EcoRI Dhet fragment complement.xdna => DNA Parallel

DNA sequence 1038 bp atgctatcaggt ... cgcgtggcttga linear

DNA sequence 12436 bp gaattctaaaca ... ctttgagaattc linear

Method: Blocks (Martinez)  
Layout: Standard  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Translation: Off

Alignment 10. Comparison of nucleotide sequence  
of SEQ ID NO:3 with the complement of the  
nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1  gaattctaaacagggcaataagggcgaaatccgaccccccatgccccgccaacacccc 60
    .                20                .                40                .                60
-----
61  ccccgatgcccccggggtctttcctggggggcctttgtcagggttgcctgtgtcacccc 120
    .                80                .                100               .                120
-----
121 acgggccccttggtcacgcttggtcgggtcaccccgagccccgggtcacgtttgccggg 180
    .                140               .                160               .                180
-----
181 gtcagccctggggcccttcgtcacgctccattgcccgctccccgcccccttgggggctt 240
    .                200               .                220               .                240
-----
241 ggctggcgcgggccggggggcgcgaccgcctctctgtccccgccccctgtcattctaccctc 300
    .                260               .                280               .                300
-----
301 caacggcgtgcagggtcccgcgccgcccccgctcccgccccaccagacacacccccctg 360
    .                320               .                340               .                360
-----
361 cgaccctgccccctcatcacgggtcacgcatggctgcgcccgcgaggcgcccccgacccc 420
    .                380               .                400               .                420
-----
421 cccgatcccccggaacgtccgcccgcaccccgcccgagccccggagccccgggacccc 480
    .                440               .                460               .                480
-----
481 cgaccccccgccgcccgcggcgtagcgccgctctgtgcgggggggctggggggcgccgcg 540
    .                500               .                520               .                540
-----
541 ggggaaggccacgccccctccactttttccaggaatgcgcggccccccatgccccacccc 600
    .                560               .                580               .                600

```

601	c g a t g c c c c c c g g g g t c t t t c c t g g g g g c c t t t g t c a g g g t t g c c t g t g t c a c c c a c	660
	. . . . . 620 . . . . . 640 . . . . . 660	
-----		
661	g g g c c c c t t g g t c a c g c t t g g c t g g g t c a c c c c g g a g c c c c g g g t c a c g t t t g c c g g g g t	720
	. . . . . 680 . . . . . 700 . . . . . 720	
-----		
721	c a g c c c t g g g g c c c t t c g t c a c g t c c a t t g c c c g t c c c c g c c c c c t t g g g g g c t t g g	780
	. . . . . 740 . . . . . 760 . . . . . 780	
-----		
781	c t g g c g c g g c c g g g g g c g c g a c c g c c t c t c t g t c c c c g c c c c c t g t c a t t c t a c c c t c c a	840
	. . . . . 800 . . . . . 820 . . . . . 840	
-----		
841	a c g g c g t g c a g g c t c c c g c g c c g c c c c c c g c t c c c g c c c c c a c c a g a c a c a c c c c c t g c g	900
	. . . . . 860 . . . . . 880 . . . . . 900	
-----		
901	a c c c t g c c c c c t c a t c a c g g t c a c g c a t g g c t g c g c c c g c g c g g c c c c c g a c c c c c c	960
	. . . . . 920 . . . . . 940 . . . . . 960	
-----		
961	c g a t c c c c g g a a c g t c c g c c g c c a t c c a c c g c c c g c a g c c c c g g a g c c c g c g g a c c c c g	1020
	. . . . . 980 . . . . . 1000 . . . . . 1020	
-----		
1021	a c c c c c c g c c g c c c g c c c g c c c g t a g c g c c g c t c t g t g c g g g g g g g c t g g g g g g c c g c g g g	1080
	. . . . . 1040 . . . . . 1060 . . . . . 1080	
-----		
1081	g g a a g g c c a c g c c c c c t c c a c t t t t t c c a g g a a t g c g c g g c c c c c c a t g c c c c a c g g g g g	1140
	. . . . . 1100 . . . . . 1120 . . . . . 1140	
-----		
1141	t c t t t c c t g g g g g g c c t t t t g t c a g g g t t g c c t g t g t c a c c c a c g g g c c c c t t g g t c a c g c	1200
	. . . . . 1160 . . . . . 1180 . . . . . 1200	
-----		
1201	t t g g c t g g g t c a c c c c g g a g c c c c g g g t c a c g t t t g c c g g g g t c a g c c c t g g g g c c c t t c	1260
	. . . . . 1220 . . . . . 1240 . . . . . 1260	
-----		
1261	g t c a c g c t c c a t t g c c c g c t c c c c g c c c c c c t t g g g g g c t t g g c t g g c g c g g c c g g g g g c	1320
	. . . . . 1280 . . . . . 1300 . . . . . 1320	
-----		
1321	g c g a c c g c c t c t c t g t c c c c g c c c c c t g t c a t t c t a c c c t c c a a c g g c g t g c a g g t c c c	1380
	. . . . . 1340 . . . . . 1360 . . . . . 1380	



[illegible]

.....

1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812

1

.....

-----

.....

1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817

-----

.....

1000 9780 9680 9560 9440 9320 9200 9080 8960 8840 8720 8600 8480 8360 8240 8120 8000 7880 7760 7640 7520 7400 7280 7160 7040 6920 6800 6680 6560 6440 6320 6200 6080 5960 5840 5720 5600 5480 5360 5240 5120 5000 4880 4760 4640 4520 4400 4280 4160 4040 3920 3800 3680 3560 3440 3320 3200 3080 2960 2840 2720 2600 2480 2360 2240 2120 2000 1880 1760 1640 1520 1400 1280 1160 1040 920 800 680 560 440 320 200 80 -20 -100 -180 -250 -320 -390 -460 -530 -600 -670 -740 -810 -880 -950 -1020 -1090 -1160 -1230 -1300 -1370 -1440 -1510 -1580 -1650 -1720 -1790 -1860 -1930 -2000 -2070 -2140 -2210 -2280 -2350 -2420 -2490 -2560 -2630 -2700 -2770 -2840 -2910 -2980 -3050 -3120 -3190 -3260 -3330 -3400 -3470 -3540 -3610 -3680 -3750 -3820 -3890 -3960 -4030 -4100 -4170 -4240 -4310 -4380 -4450 -4520 -4590 -4660 -4730 -4800 -4870 -4940 -5010 -5080 -5150 -5220 -5290 -5360 -5430 -5500 -5570 -5640 -5710 -5780 -5850 -5920 -5990 -6060 -6130 -6200 -6270 -6340 -6410 -6480 -6550 -6620 -6690 -6760 -6830 -6900 -6970 -7040 -7110 -7180 -7250 -7320 -7390 -7460 -7530 -7600 -7670 -7740 -7810 -7880 -7950 -8020 -8090 -8160 -8230 -8300 -8370 -8440 -8510 -8580 -8650 -8720 -8790 -8860 -8930 -9000 -9070 -9140 -9210 -9280 -9350 -9420 -9490 -9560 -9630 -9700 -9770 -9840 -9910 -9980 -10000

End use of the data is not controlled by the data controller. The data controller is not responsible for the use of the data by the data user.

2101 ccgtagcgccgctctgtgcggggggctggggggccgcgggggaaggccacgccccctcc 2160  
 . 2120 . 2140 . 2160



2941 gttatgagtgactggactggaggagccctccttgctcctctattcctttgctctcatgctt 3000  
 . 2960 . 2980 . 3000

3001 ataattataat<sup>ttt</sup>gatcatct<sup>ttt</sup>atct<sup>ttc</sup>agaagagac<sup>ttt</sup>ctctgtccacttggagcc 3060

3061 ctttgtatactcctactgatgagtaagtattacacctttgccccacacccccctttccct 3120  
 . 3080 . 3100 . 3120

```

      .               20
1  ---atgctatcaggtaacgcagg----- 20
      xxxxxxxxxxxx|||||xx
3121 tactcttctctctctctaagcactttctctcttccccagtcaccctcctgctcatcgct 3180
      .               .               .               .
      3140             3160             3180

```

3181 ctctggaatttgcacggacaggcattgttccttggaattgtgctgttcattcttcgggtgc 3240  
 . 3200 . 3220 . 3240

3241 ttacttggtaagatctaacattccctaggaattattttaccacacccccactttttccaacc 3300  
 . 3260 . 3280 . 3300

3301 ctaacactcttttttcaacgcagtccttaggtatctggatctacttattggagatgctctg 3360  
 . 3320 . 3340 . 3360

3361 gcgacttggtgccaccatctgtggcagcttttggccttcttctctagccttcttctctagacct 3420  
 . 3380 . 3400 .. 3420

3421 catcctgctcattattgctctctatctacaacaaaactgggtggactctattggttgatct 3480  
 . 3440 . 3460 . 3480

3481 cctttggctcctcctgtttctggcgattttaatctggatgtattaccatggacaacgaca 3540  
 . 3500 . 3520 . 3540

3541 cagtgatgaacaccaccacgatgactccctcccgccacctcaacaagctaccgatgattc 3600  
 . 3560 . 3580 . 3600

3601 tggccatgaatctgactctaactccaacgagggcagacaccacctgctcgtgagtggagc 3660  
 . 3620 . 3640 . 3660

3661 cggcgacggaacccactctgctctcaaacctaggcgacacctggaggtggctctgacaa 3720  
 . 3680 . 3700 . 3720

3721 tggccacaggaccctgacaacactgatgacaatggccacaggaccctgacaacactga 3780  
 . 3740 . 3760 . 3780

3781 tgacaatggcccatgacccgctgcctcaggaccctgacaacactgatgacaatggccc 3840  
 . 3800 . 3820 . 3840

3841 acaggaccctgacaacactgatgacaatggcccacatgaccgctgcctcatagccctag 3900  
 . 3860 . 3880 . 3900

3901 cgactctgctggaaatgatggaggccctccacaattgacggaagaggttgaaaacaaagg 3960  
 . . . . . 3920 . . . . . 3940 . . . . . 3960

3961 aggtgaccaggggccgcctttgatgacagacggaggcgggcggtcatagtcatttcggg 4020  
 . 3980 . 4000 . 4020

4021 ccatggcggcggtgatccacaccttcctacgctgcttttggggttcttctgggttcgggtgg 4080  
 . 4040 . 4060 . 4080

4081 agatgatgacgacccccacggcccagttcagctaagctactatgactaacctttctttac 4140  
 . 4100 . 4120 . 4140

4141 ttctaggcattaccatgtcataggcttgcctgactgactctccctccatttactgggaat 4200  
 . 4160 . 4180 . 4200

4201 gccttagctaatacaccttaactggcacacactcccttagccacactgtctgtctaggctg 4260  
 . 4220 . 4240 . 4260

4261 aaaagccacattcatatttctattttcaaaacaaggggaaaggaggacatgcgagaattggc 4320  
 . 4280 . 4300 . 4320

4321 agacacctttaccagcccttaacacaccacacaggtagcaaggaccggcggttgccag 4380  
 . 4340 . 4360 . 4380

4381 actccgccaccaacgcccctgcgttgaaacccacccctcctacacacatcagacctctgca 4440  
 . 4400 . 4420 . 4440

4441 caacacaactaccaggcagatgaggcccttacttccacaggggtactggcataccagcgg 4500  
 . 4460 . 4480 . 4500

---

---

.....

-----

-----

.....

-----

---

21 -----agaaggaagc----- 29

-----

-----

.....

5221 cagcgcgtgggggtgtcaacaaagaactttgacctgttgccctgagatgtgaattgggat 5280  
 . 5240 . 5260 . 5280

5281 ggtgtggataaacatctcccgctagatggcgcccttattattgatgtgacttgtgatgcaa 5340  
 . 5300 . 5320 . 5340

5341 taaataaaaagtacagatagatggcactcttaccttctctgcccgccttcttcgtatatgt 5400  
 . . . . . 5360 . . . . . 5380 . . . . . 5400

5401 gttgagatgagtcacccgtggagagtagggagggggagggagcccgtcattcccgtcgt 5460  
 . 5420 . 5440 . 5460

5461 gttgcaatcccaagtacagactttgatcttgggttcttagtggttgatagtcgcagagtgc 5520  
 . 5480 . 5500 . 5520

5521 ggtcgccattgccccaatatgggtcctcataaggcggtgggggctcttcattagattcac 5580  
 . 5540 . 5560 . 5580

5581 gttcctcatcgttcgggtgggggtgggggtgttcccagaagagccagaagcagatggatatt 5640  
 . 5600 . 5620 . 5640

5641 gggagttgtttccgccatcgtaaacatccggatccccgcggggctagggggaccgcgc 5700  
 . 5660 . 5680 . 5700

5701 ccattggcaccattttctagggaccccatagctgcagcagcgactgcaaaaccggcctgaa 5760  
 . 5720 . 5740 . 5760

5761 atgctctctgagaaaacaaggcgagagggtattacgccccagcaagcttcaggcaaaaataa 5820  
 . 5780 . 5800 . 5820

5821 cctgtgatgctgaaaccgcggcgctacaggatcaccccaattggcaagacctgggcggag 5880  
 . 5840 . 5860 . 5880

5881 ttcctgagagcccaggggtctcgtgcagggtgtccccggggaattgttctccctgatcacc 5940  
 . 5900 . 5920 . 5940

5941 gccacccccgttttctccaaaccagcagctgtgacaatottcacacactgctgctgtca 6000  
 . 5960 . 5980 . 6000

6001 cctggaactatatttcccacggtgcccttccgccattttcccacgagtcgcgaggctatc 6060  
 . 6020 . 6040 . 6060

```

-----
6061 caccgcgaatgccaccccccaatgccacactaaaacaagggtgaaataggcaagtgcgtt 6120
      .               .               .               .
      6080               6100               6120
-----
6121 tattgcgacaagtatccagaaacataaacccccgtgggcttcctccttgtcatttttccca 6180
      .               .               .               .
      6140               6160               6180
30 -----aacag----- 34
      |||||
6181 acgcagggtcactggcagggtgccagggttgggaagtgcagggtcaacagcaacagagagg 6240
      .               .               .               .
      6200               6220               6240
-----
6241 ctcccatccttttcttcataacaccgccatttgccgcagttggtgcgggctccacgccc 6300
      .               .               .               .
      6260               6280               6300
-----
6301 tcgggcatgagccactggacgtggggatggggaaatgcattcacgggtgcatgtcacagta 6360
      .               .               .               .
      6320               6340               6360
-----
6361 aggacagagaagtctgggaactgagaccttctcgagtgagacagacagcgtttagaggcttc 6420
      .               .               .               .
      6380               6400               6420
-----
6421 accacgctcagggtgttcctgcttggtagacctcggtctcgccagtttcatgcggcacagg 6480
      .               .               .               .
      6440               6460               6480
-----
6481 tagttgccgtcatgggagatggttgccagcggtgactactaaaaagaagggtgttgccactt 6540
      .               .               .               .
      6500               6520               6540
-----
6541 ctgtggatatcaaagaagcccctgaaaggccactctataaagatgacatcggtggtgcatg 6600
      .               .               .               .
      6560               6580               6600
-----
6601 cgcccaataagcacctgctcctctcctggggcccagtttaaacagctgacctcaatctct 6660
      .               .               .               .
      6620               6640               6660
-----
6661 ggaccgagggtcacccctcctccagtaggaggtcagggtgactcgctcacccaagaaagcg 6720
      .               .               .               .
      6680               6700               6720
35 -----cctg----- 38
      |||||
6721 gtgacagcctggcgggcgccacacaggaggccaacaggaggagctgagcgatgaacctg 6780
      .               .               .               .
      6740               6760               6780
-----
6781 gccattgctctggactctcctcaccagggcctcgcggtcttatactattctgccacgccc 6840
      .               .               .               .
      6800               6820               6840

```

6841 ttttatcatataagcctgaagcccgtagctggcctgacgagaccatgaggccagccaag 6900  
 . 6860 . 6880 . 6900

6901 tctacagattctgtgtttgtgaggacccccggtcgaggcggtgggtcgcgccctcgccgcgcg 6960  
 . 6920 . 6940 . 6960

6961 gacgacaaggtggctgagtcacgctacctcatgttcagggccatgtacgcggtgttcacc 7020  
 . 6980 . 7000 . 7020

7021 cgggatgagaaagacctgcctttgccagccctggctctctgccggctcatcaaggcctcc 7080  
 . 7040 . 7060 . 7080

39 -----cggaggt 45  
|||  
7081 ctgaggaaggataggaagctgtacgcggagctggcctgcaggacagccgacatcgggggc 7140  
.                 7100                 7120                 7140

46 tcggccgcccgcgggccaggacctcatcagcgctccccgcgaacacctttatgaca----- 99  
| | | | | | | | | | | | | | | | | | | | | |  
7141 aaagacacgcacgtacggctcatcagcgctcctgcgcgcagtgtacaacgaccactac 7200  
          :                 7160                 :                 7180                 :                 7200

7201 gactactggtcgcggtctcaggggtggtgctgtgctacacagtgggtgtttgcgggtgcgaac 7260  
 . 7220 . 7240 . 7260

7261 tacctggatgaccacaagagcgccgccttcgtgctgggggcaatcgcccactacctggcc 7320  
 . 7280 . 7300 . 7320

7321 ctctatcgcagactctgggtttgcgaggctgggcggcatgccaaagatcgctgagacgtcag 7380  
 . 7340 . 7360 . 7380

7381 ttccccgtgacgtggggccctggccagcctgactgacttctgaaatctttgtaaatgaat 7440  
 . 7400 . 7420 . 7440

7441 aaacagtgggtgttgcgatgagtaaaagtgtaacatttaattgtgggactgggaggccgg 7500  
 . 7460 . 7480 . 7500

7501 ggcgataccttgggcatcatgcagggtgcacagactagcgaggataatctgggcagccag 7560  
 . 7520 . 7540 . 7560

7561 agccagccgggtccgtgcggctacatctacttttaccacctggccacctacctcttagg 7620  
 . 7580 . 7600 . 7620



-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

8341 agcgcggcatgcgagctggcggtgagctacgcacccgccatgctggaggcctcgcacgag 8400  
 . 8360 . 8380 . 8400

9121 ctggagcggctgtctcggctgggcatagacgctgaaggaaagctagagaaggagccccat 9180  
 . 9140 . 9160 . 9180

9181 ggccgcgctgactttgtcaagatgttcaaggacgtggatgcggcggtggacgccgaagtg 9240  
 . 9200 . 9220 . 9240

9241 gtccagtttatgaacagcatggccaagaacaacatcacctacaaggacctgggtcaagagc 9300  
 . 9260 . 9280 . 9300

9301 tgctaccacgtgatgcagtactcgtgcaacccctttgcgcagcccgccctgccccatcttc 9360  
 . 9320 . 9340 . 9360

9361 acccagctgtttttaccgctcactgctgaccatcctgcaggacatctccctgcccatctgt 9420  
 . 9380 . 9400 . 9420

9421 atgtgctatgagaatgacaaccccgggcttgccagagccccccagagtggctaaaggggt 9480  
 . 9440 . 9460 . 9480

9481 cactaccagacgctgtgaccaaacttttaggagcctggccatcgacaaggggggtcctcacg 9540  
 9500 9520 9540

9541 gccaaggaggccaaggtggtgcatggggagcccacctgcgacctgccagacctggacgcg 9600  
 . 9560 . 9580 . 9600

9601 gccctgcagggccgggtgtacggccggcggtgcctgtgcgcatgtccaaggtgctgatg 9660  
 . 9620 . 9640 . 9660

9661 ctgtgccccaggaacatcaagatcaagaacaggggtggtcttcacgggggagaatgccgcc 9720  
 . 9680 . 9700 . 9720

9721 ctccagaacagcttcatcaagtccactaccaggagggagaactacatcatcaacggggccc 9780  
 . 9740 . 9760 . 9780

9781 tacatgaaattcctcaacacctaccacaagaccctattcccggaactaagctctcaagc 9840  
 . 9800 . 9820 . 9840

9841 ctgtacctgtggcacaacttttccaggcggcgctcgggtccctgtccccagcggggccagc 9900  
 . 9860 . 9880 . 9900

9901 gcggaggagtactctgacctggccctctttgtggacgggggctcccggggccacgaagag 9960  
 . . . . .  
 9920 9940 9960

```

-----
9961 agcaacgtcatagatgtggtgcctggcaacctggtcacttacgccaaagcagagggtcaac 10020
      .               .               .               .
      9980               10000               10020
-----

10021 aacgccatcctgaaggcgtgcgccagaccagttctacatcagcctgattcagggactg 10080
      .               .               .               .
      10040               10060               10080
-----

10081 gtgccgaggacgcagtcggtgcccggccgtgactacccccacgtactgggcacgcggggcg 10140
      .               .               .               .
      10100               10120               10140

149 -----cctacgcg----- 157
      |||||
10141 gtggagtcggcagcggcctacgcgaggccacctcctcccttactgcgaccacggtggtc 10200
      .               .               .               .
      10160               10180               10200

158 -----ccccgctgccc 168
      | |||||x
10201 tgcgcgccacagactgtcttagccaggtctgcaaggcccggttgcacgctgcca 10260
      .               .               .               .
      10220               10240               10260

169 cccctttccaccagg----- 184
xxxxxxxxxxxxxxxx
10261 gtgaccatcaacaagtacacgggggtcaacggcaacaaccagatattccaggccgggaac 10320
      .               .               .               .
      10280               10300               10320
-----

10321 ctgggatactttatgggcccggcggtggacaggaacctgctgcaggcccccggggctggg 10380
      .               .               .               .
      10340               10360               10380
-----

10381 ctgcgcaagcaggccgggggtcttccatgcggaagaagtttgtctttgccacccccacc 10440
      .               .               .               .
      10400               10420               10440
-----

10441 ctaggggttgaccgtgaagcgccggacccaagccgcgaccacatatgagattgagaacatc 10500
      .               .               .               .
      10460               10480               10500
-----

10501 agggctggcctggaggccattatatcacaaaaacaggaggaagactgtgtgtttgatgtg 10560
      .               .               .               .
      10520               10540               10560
-----

10561 gtgtgcaaccttgtggatgccatgggcgaggcatgcgcctcgctgactagggacgacgcg 10620
      .               .               .               .
      10580               10600               10620
-----

10621 gagtacttattgggcccgttctccgtcctggcggacagcgtcctagaaacctggcgacc 10680
      .               .               .               .
      10640               10660               10680
-----

10681 attgcctccagcgggatagagtggacggcgaggccgctcgggactttctggaggaggagtg 10740
      .               .               .               .
      10700               10720               10740

```

10741 tgggggtgggcccggggcagcccaggaactttatcagcgtggccgagccggtcagcacc 10800  
 . 10760 . 10780 . 10800

10801 gcgtcgcaggcctcgccgggctgctgctgggtggaggaggccagggtccgggggcaga 10860  
 . 10820 . 10840 . 10860

10861 cgcaagcgccgtctggccaccgttctccccggactcgaggtctagagaccctggggcg 10920  
 . 10880 . 10900 . 10920

10921 cgatgtcggggctgctggcggcgcgctacagccagggtgtacgccctggcggttgagctga 10980  
 . 10940 . 10960 . 10980

10981 gcgtgtgcacccggctggacccccggagtctggacgtggctgcggtggtgcgcaacgccg 11040  
 . 11000 . 11020 . 11040

11041 gcctgctggccgagctggaggccatcctccttccccgtttgagacggcagaatgaccgtg 11100  
 . 11060 . 11080 . 11100

11101 catgcagcgcacctgtccctggagctggtgcacctgctagagaactcgagagaggcctctg 11160  
 . 11120 . 11140 . 11160

11161 ccgcgctgctcgcccctggtagaaaagggtacccgggtcccgccctctccgtacccctcag 11220  
 . 11180 . 11200 . 11220

11221 tcgcgtactctgtggagttttacggggggcataaagtcgatgtaagtttgtgcctaataa 11280  
 . 11240 . 11260 . 11280

185 -----caatagc----- 191

11281 atgacatagagatTTTaatgaagagaatcaatagcgtgttttattgcatgtctcacacca 11340  
 . 11300 . 11320 . 11340

11341 tggggctggagagcctggaacgggccctggatctgctgggccgctttcggggcgtaagtc 11400  
 . 11360 . 11380 . 11400

11401 ccattcccagaccgcgcctctacatcacctctgtgcccctgctggcgctgtgtgggcgagc 11460  
 . 11420 . 11440 . 11460

11461 tgatggttctgcccaaccacggcaacccttcacggcagaggggacccacgtctcctgta 11520  
 . 11480 . 11500 . 11520

11521 accacctggcgggtgccgggtgaatccggagccgggtctcgggactgtttgagaatgaagtcc 11580  
 . 11540 . 11560 . 11580

11581 gccaggcggggctcgggcacctgttgaggctgaggagaaggcgaggccgggcggccag 11640  
 . 11600 . 11620 . 11640

192 --caccgcgccttctacgggtcctgg-ggccgag----- 223  
| | | | | | | | | |  
11641 aggaggcgcggtcccgggcccggggcgccgagggcagagggggcgaccagagcgctgg 11700  
. 11660 . 11680 . 11700

11701 acacctacaacgtcttctcgcacagtgcggcgagggtggcgaggctctcggagctcctct 11760  
 . 11720 . 11740 . 11760

11761 attggaactctggcgcccatgctatcggtgcaacggggcaggggaggggtggcgccatt 11820  
 . 11780 . 11800 . 11820

11821 cccgcctctctgcctgtttgcccgaggcgtcgcctggccctggtgcgggggcctgcg 11880  
 . 11840 . 11860 . 11880

11881 aggagggcgtggcgggggcaaggctgactcacctgtttgacgccgtggctcccggggcca 11940  
 . 11900 . 11920 . 11940

224 -----cgg-tc----- 228  
         ||| ||  
11941 cggagcggtcttctgcggcggtctacagctcctcgggcgacgcggtggaggcgctga 12000  
               ·                    ·                    ·                    ·  
                 11960                 11800                 12000

12001 aggcggactgcgccgccgccttcacggcgacccccagtaccgggccatcctgcaaaaga 12060  
 . 12020 . 12040 . 12060

12061 ggaacgagctgtacacgcggctcaaccgagccatgcagcggttgggccgaggcgaggagg 12120  
 . 12080 . 12100 . 12120

229 -----gccccggccggcggtactttacctccccaggagg 263

12121 aggcgtcccgaggagagcccggaggtgccccggccggc===== 12157  
          .  
               12140                .

264            .                 280                 .                 300                 .                 320  
264 ttactacgccgggccccgcggggcggggaaccgggtgccttcttggcgatggacgctcacac 323

324 ctaccacccccacccacacccccctccggcctactttgggttgccgggctcttttggccc 383

12383 ctcgggtccctgggaggacaggcgcgccggcgaggggtgccgcctttgagaattc 12436  
 . 12400 . 12420 .

```
% Identity = 1.8 (239/13074)
```

9310-13DVCTDV SEQ ID NO 3.xdna x Bankier et al. EcoRI Dhet fragment/18/07/5-02/4/PM=> DNA Parallel

///



9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BA-LF3.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ\*

Protein sequence 609 aa \*GRRGVLIGPLL ... DRRAGEGAAAFEN

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 11.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BA-LF3, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .           40           .           60
1  MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPNDVVFREAQRSYLVFLTSQFCYEEY 60
      RR           R           P +P           Q           +           Q Y+
1  *GRRGVLIGPLLRRPGGQRPRNPGDHCLQRDRVDGGGRSGLSGGSVGVWARGSPGQL=YQRG 59
      .           20           .           40           .           60
61 VQRTFGVPRRQRAIDKRQRASVAGAGAHHLGGSSATPVQQAQAAASAGTGALASSAPST 120
      V           RA           +RA + G           S T           + AA +G A A S
60 RAGQHRVAGLGRAGAAAGWRRAGLRGQTQAPSGHRSPRTRGLETPGAAMSGLLAAAYSQVYA 119
60 .           80           .           100           .           120
121 -AVAQSATPSVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ*-- 177
      AV S           +           + A           A A           A +           +A
120 LAVELSVCTRLDPRSLDVAAVVRNAGLLAELEAILLPRLRRQNDRACSALSLELVHLLN 179
120 .           140           .           160           .           180
-----
180 SREASAALLAPGRKGTRVPPLRTPSVAYSVEFYGGHKVDVSLCLINDIEILMKRINSVFY 239
180 .           200           .           220           .           240
-----
240 CMSHTMGLESLEALRDLGLRFRGVSPIDPRLYITSVPCWRCVGLMVLNHNPNSTAEG 299
240 .           260           .           280           .           300
-----
300 THVSCNHLAVPVNPEPVSGLFENEVRQAGLGHLLAEAEKARPGGPPEGAVPGPGRPEAEG 359
300 .           320           .           340           .           360
-----
360 ATRALDTYNVVFSTVPPEVAELSELLYWNSSGGHAIGATGQGEAGGHSRLSALFARERRLAL 419
360 .           380           .           400           .           420
-----
420 VRGACEEALAGARLTHLFDVAVPGATERLFCGGVYSSSGDAVEALKADCAAFTAHPQYR 479
420 .           440           .           460           .           480
-----
480 AILQKRNELYTRLNRAMQRLGRGEEEEASRESPEVPRPAGAREPGPSGALSALKRKEQYL 539
480 .           500           .           520           .           540
-----
540 RQVATEGLAKLQSCLAQQSETLTETLCLRVWGDVVYWELARMRNHFLYRRAFVSGPWEDR 599
540 .           560           .           580           .           600

```

-----  
600 RAGEGAAFEN  
600

609

% Identity = 4.9 (30/610)      % Homology = 2.0 (12/610)      % Total = 6.9 (42/610)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BA-LF3.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA\*

Protein sequence 609 aa \*GRRGVLIGPLL ... DRRAGEGAAFEN

```
Method:           Diagonals (BLOSUM62)
Layout:           Standard
Block Length ≤:   6-aa
Mismatch penalty:  Smaller (1)
Gap penalty:      Medium (2)
Display:          BLOSUM62
```

**Alignment 12.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BA-LF3, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

[illegible]

-----  
601 AGEGAAFEN

609

% Identity = 9.7 (59/609)

% Homology = 3.0 (18/609)

% Total = 12.6 (77/609)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BA-LF3.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein sequence 609 aa \*GRRGVLLIGPLL ... DRRAGEGAAFEN

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

Alignment 13. Comparison of the amino acid sequence of  
SEQ ID NO:5 with the amino acid sequence, BA-LF3, encoded  
by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1 *GRRGVLLIGPLLPGGQRP RNP GDHCLQRDRVDGGGRSGLSGGSVGVWARGSPGQLYQRGR 60
      .               20               .               40               .               60
-----
61 AGQHRVAGLGRAAAGWRRAGLRGQTQAPSGHRSPRTRGLETPGAAMSGLLAAAYSQVYAL 120
      .               80               .               100              .               120
-----
121 AVELSVCTRLDPRSLDVAAVVRNAGLLAELEAILLPRLRRQNDRA CSALSLELVHLLENS 180
      .               140              .               160              .               180
-----
181 REASAALLAPGRKGTRVPPLRTPSVAYSVEFYGGHKVDVSLCLINDIEILMKRINSVFYC 240
      .               200              .               220              .               240
-----
241 MSHTMGLESLE RALDLLGRFRGVSPIDPRLYITSVPCWRCV GELMVLPNHGNPSTAEGT 300
      .               260              .               280              .               300
-----
301 HVSCNHLAVPVNPEPVSGLFENEVRQAGLGHLLEAE EKARPGGP EEGAVPGPGRPEAEGA 360
      .               320              .               340              .               360
-----
361 TRALDTYNVFSTVPPEVAELSELLYWNSSGGAIGATGQGE GGGHSRLSALFARERRLALV 420
      .               380              .               400              .               420
-----
421 RGACEEALAGARLTHLFD AVAPGATERLFCGGVYSSSGDAVEALKADCAA AFTAHPQYRA 480
      .               440              .               460              .               480
-----
481 ILQKRNELYTRLN RAMQRLGRGEEEEASRESPEVPRPAGAREPGPSGALS DALKRKEQYLR 540
      .               500              .               520              .               540
1 -----AVDTGSGGGGQP HDT 15
      .               15
      G
541 QVATEGLAKLQSC LAQQSETLTETLCLRVWGDVVYWELARMRNHFLYRR AFVSGPWEDRR 600
      .               560              .               580              .               600

```

20  
16 APRGARKKQ  
A GA +  
601 AGEGAAFEN

24

609

% Identity = 0.7 (4/609)      % Homology = 0.2 (1/609)      % Total = 0.8 (5/609)

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BA-LF3.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 609 aa \*GRRGVLLIGPLL ... DRRAGEGAAFEN

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

Alignment 14. Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BA-LF3, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1 *GRRGVLLIGPLL R PGGQRPRNPGDHCLQRDRVDGGGRSGLSGGSVGWARGSPGQLYQRGR 60
   .                20                .                40                .                60
-----
61 AGQHRVAGLGRAAAGWRRAGLRGQTQAPSGHRSPRTRGLETPGAAMSGLLAAAYSQVYAL 120
   .                80                .                100               .                120
-----
121 AVELSVCTRLDPRSLDVAAVVRNAGLLAELEAILLPRLRRQNDRACSALELVHLLENS 180
   .                140               .                160               .                180
-----
181 REASAALLAPGRKGTRVPPLRTPSVAYSVEFYGGHKVDVSLCLINDIEILMKRINSVFYC 240
   .                200               .                220               .                240
-----
241 MSHTMGLESLERALDLLGRFRGVSPIDPRLYITSVPCWRCV GELMVLPNHGNPSTAEGT 300
   .                260               .                280               .                300
-----
301 HVSCNHLAVPVNPEPVSGLFENEVRQAGLGHLLEAE EKARPGGP EEGAVPGPGRPEAEGA 360
   .                320               .                340               .                360
-----
361 TRALDTYNVFSTVPPEVAELSEL LYWNSSGGHAIGATGQGE GGGHSRLSALFARERRLALV 420
   .                380               .                400               .                420
-----
421 RGACEEALAGARLTHLFDVAPGATERLFCGGVYSSSSG DAVEALKADCAAFTAHPQYRA 480
   .                440               .                460               .                480
-----
481 ILQKRNELYTRLN RAMQRLGRGEEEEASRESPEVPRPAGAREPGPSGALS DALKRKEQYLR 540
   .                500               .                520               .                540
1 -----STAVAQSATPSVSSSI SLRA 21
   .                500               .                520               .                540
541 QVATEG LAKLQSC LAQQSETLTETLCLRVWGDVVY WELARMRNHFLYRRAFVSGPWEDRR 600
   .                560               .                580               .                600

```

22 ATSGATAAA  
A GA  
601 AGEGAAFEN

30

609

% Identity = 0.8 (5/609)

% Homology = 0.2 (1/609)

% Total = 1.0 (6/609)

///



9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BA-LF2.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ\*

Protein sequence 1129 aa MQGAQTSEDNLG ... RLATVLPGLEV\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 15.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BA-LF2, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1  MQGAQTSEDNLG SQSQPGPCGYIIFYPLATYPLREVATLGTGYAGHRCLTVPLLCGITVE 60
   .              20              .              40              .              60
-----
61 PGFSINVKALHRRPDPNCGLLRATSYHRDIYVFHNAHMPPIFEGPGLEALCGETREVF 120
   .              80              .              100             .              120
-----
121 YDAYSALPRESSKPGDFFPEGLDPSAYLGAVAITEAFKERLYSGNLVAIPSLKQEVAVGQ 180
   .              140             .              160             .              180
-----
181 SASVRVPLYDKEVFPEGVPQLRQFYNSDLRCMHEALYTGLAQALRVRVVGKLVELLEKQ 240
   .              200             .              220             .              240
-----
241 SLQDQAKVAKVAPLKEFPASTISHPDGALMIVDSAACELAVSYAPAMLEASHETPASLN 300
   .              260             .              280             .              300
-----
301 YDSWPLFADCEGPEARVAALHRYNASLAPHVSTQIFATNSVLYVSGVSKSTGQGKESLFN 360
   .              320             .              340             .              360
-----
361 SFYMTHTGLGTLQEGTWDP CRRPCFSGWGGPDVTGTNGPGNYAVEHLVYAASFSPNLLARY 420
   .              380             .              400             .              420
-----
421 AYYLQFCQGQKSSLTVPVETGSYVAGAAASPMCSLCEGRAPAVCLNTLFFRLRDRFPPVM 480
   .              440             .              460             .              480
-----
481 STQRRDPYVISGASGSYNETDFLGNFLNFIDKEDDGQRPDDEPRYTYWQLNQNLLERLSR 540
   .              500             .              520             .              540
-----
541 LGIDAEGKLEKEPHGPRDFVKMFKDVAADVAE VQFMNSMAKNNITYKDLVKSCYHVMQ 600
   .              560             .              580             .              600

```

601 YSCNPFAQPA CFIFTQLFYRSLLTILQDISLPICMCYENDNPGLGQSPPEWLKGHYQTLC 660  
 . 620 . 640 . 660

661 TNFRSLAIDKGVLTAKVVGHEPTCDLPDLDAALQGRVYGRRLPVRMSKVLMLCPRNI 720  
 . 680 . 700 . 720

721 KIKNRVVFTGENAALQNSFIKSTTRRENYIINGPYMKFLNTYHKTLFPDTKLSSLYLWHN 780  
 . 740 . 760 . 780

781 · FSRRRSVPVPSGASAE EYSDLALFVDGG SRAHEESNVIDVVPGNLV TYAKQRLNNA ILKA 840  
· 800 · 820 · 840

841 CGQTQFYISLIQGLVPRTQSVPARDYPHVLGTRAVESAAAYAEATSSLTATTVVCAATDC 900  
 . 860 . 880 . 900

1 -----MARRLPKP 8

901 LSQVCKARPVVTLPTINKYTG VNGNNQIFQAGNLGYFMGRGVDRNLLQAPGAGLRKQAG 960

• 20 • 40 • 60

9 TLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEYVQRTFGVP 6

9 TLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEYVQRTFGVP 68  
 xxx  
 961 GSSMRKKFVFATPTLGLTVKRRTQAATTYEIENIRAGLEAIISQKQEEDCVFDVVCNLVD 1020  
 . 980 . 1000 . 1020

80	100	120
----	-----	-----

69 RRQRAIDKRQRSVAGAGAAHLGGSSATPVQQAQAAAASAGTGALASSAPSTAVAQSATP 128  
XX

1 AMGEACASLTRDDAEYLLGRFSVLADSVLETLATIASSGIEWTAEAARDFLEGVWGGPGA 1

.	140	.	160	.	180
---	-----	---	-----	---	-----

19 SVSSSISSLRAATSGATAAASAAA VDTGSGGGGQPHDTAPRGARKKQ\*

```

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
1081 AQDNFISVAEPVSTASQASAGLLLGGGGQSGGRRKRRLATVLPGLEV*      1129
      .              .              .              .
      1100              1120

```

% Identity = 0.0 (0/1129)      % Homology = 0.0 (0/1129)      % Total = 0.0 (0/1129)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BA-LF2.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA\*

Protein sequence 1129 aa MQGAQTSEDNLG ... RLATVLPGLEV\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 16.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BA-LF2, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1  MQGAQTSEDNLGSQSQPGPCGYIYFYPLATYPLREVATLGTGYAGHRCLTVPLLCGITVE 60
   .                20                .                40                .                60
-----
61 PGFSINVKALHRRPDPNCGLLRATSYHRDIYVFHNAHNVPPPIFEGPGLEALCGETREVFG 120
   .                80                .                100               .                120
-----
121 YDAYSALPRESSKPGDFFPEGLDPSAYLGAVAITEAFKERLYSGNLVAIPSLKQEVAVGQ 180
   .                140               .                160               .                180
-----
181 SASVRVPLYDKEVFPEGVPQLRQFYNSDLSRCMHEALYTGLAQALRVRRVGKLVELLEKQ 240
   .                200               .                220               .                240
-----
241 SLQDQAKVAKVAPLKEFPASTISHPDSGALMIVDSAACELAVSYAPAMLEASHETPASLN 300
   .                260               .                280               .                300
-----
301 YDSWPLFADCEGPEARVAALHRYNASLAPHVSTQIFATNSVLYVSGVSKSTGQGKESLFN 360
   .                320               .                340               .                360
-----
361 SFYMTHTGLGTLQEGTWDPCCRPCFSGWGGPDVTGTNGPGNYAVEHLVYAASFSPNLLARY 420
   .                380               .                400               .                420
-----
421 AYYLQFCQGQKSSLTPVPETGSYVAGAAASPMCSLCEGRAPAVCLNTLFFRLRDRFPPVM 480
   .                440               .                460               .                480
-----
481 STQRRDPYVISGASGSYNETDFLGNFLNFIDKEDDGQRPDDEPRYTYWQLNQNLRLSLR 540
   .                500               .                520               .                540
-----
541 LGIDAEGKLEKEPHGPRDFVKMFKDVDAAVDAEVVQFMNSMAKNNITYKDLVKSCYHVMQ 600
   .                560               .                580               .                600

```

```

-----
601 YSCNPFAQPACPIFTQLFYRSLTILQDISLPICMCYENDNPGLGQSPPEWLKGHYQTL 660
      620                      640                      660
-----

661 TNFRSLAIDKGVLTAKVHGEPTCDLPDLDAALQGRVYGRRLPVRMSKVLMLCPRNI 720
      680                      700                      720
-----

721 KIKNRVVFTGENAALQNSFIKSTTRRENIINGPYMKFLNTYHKTLPDTKLSSLYLWHN 780
      740                      760                      780
      20                      40
1 ---MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPP 57
      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
781 FSRRRSVPVPSGASAEYSIDLALFVDGGSRAHEESNVIDVPGNLVTYAKQRLNNAILKA 840
      800                      820                      840
      60                      80                      100
58 FSHQAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPPPP 117
      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
841 CGQTQFYISLIQGLVPRTQSVPARDYPHVLGTRAVESAAAYAEATSSLTATTVVCAATDC 900
      860                      880                      900
      120                      140                      160
118 AYFGLPGLFGPPPPCLLTDSHLRADYVPAPSRSNKRKRDPEDDEEGGLFPGEDATLYR 177
      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
901 LSQVCKARPVVTLPVTINKYTG VNGNNQIFQAGNLGYFMGRGVDRNLLQAPGAGLRKQAG 960
      920                      940                      960
      180                      200                      220
178 KDIAGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSIEVPR 237
      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
961 GSSMRKKFVFATPTLGLTVKRRTQAATTYEIENIRAGLEAIIISQKEEDCVFDVVCNLVD 1020
      980                      1000                      1020
      240                      260                      280
238 YVPHPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTE 297
      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
1021 AMGEACASLTRDDAEYLLGRFSVLADSVLETLATIASSGIEWTAEAAARDFLEGVWGGPGA 1080
      1040                      1060                      1080
      300                      320                      340
298 GHHRGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA* 346
      XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
1081 AQDNFISVAEPVSTASQASAGLLGGGGQSGGRRKRRLATVLPGLEV* 1129
      1100                      1120

```

% Identity = 0.0 (0/1129)      % Homology = 0.0 (0/1129)      % Total = 0.0 (0/1129)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BA-LF2.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein sequence 1129 aa MQGAQTSEDNLG ... RLATVLPGLEV\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 17.** Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BA-LF2, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1 MQGAQTSEDNLGSQSQPGPCGYIIFYPLATYPLREVATLGTGYAGHRCLTVPLLCGITVE 60
  .                20                .                40                .                60
-----
61 PGFSINVKALHRRDPNCGLLRATSYHRDIYVFHNAHMPPIFEGPGLEALCGETREVFG 120
  .                80                .                100               .                120
-----
121 YDAYSALPRESSKPGDFFPEGLDPSAYLGAVAITEAFKERLYSGNLVAIPSLKQEVAVGQ 180
  .                140               .                160               .                180
-----
181 SASVRVPLYDKEVFPEGVPQLRQFYNSDLSRCMHEALYTGLAQALRVRVVGKLVELLEKQ 240
  .                200               .                220               .                240
-----
241 SLQDQAKVAKVAPLKEFPASTISHPDGALMIVDSAACELAVSYAPAMLEASHETPASLN 300
  .                260               .                280               .                300
-----
301 YDSWPLFADCEGPEARVAALHRYNASLAPHVSTQIFATNSVLYVSGVSKSTGQGKESLFN 360
  .                320               .                340               .                360
-----
361 SFYMTHTGLGTLQEGTWDPFCRRPCFSGWGGPDVTGTNGPGNYAVEHLVYAASFSPNLLARY 420
  .                380               .                400               .                420
-----
421 AYYLQFCQGQKSSLTPVPETGSYVAGAAASPMCSLCEGRAPAVCLNTLFFRLRDRFPPVM 480
  .                440               .                460               .                480
-----
481 STQRRDPYVISGASGSYNETDFLGNFLNFIDKEDDGQRPDDEPRYTYWQLNQNLRLSR 540
  .                500               .                520               .                540
-----
541 LGIDAEGKLEKEPHGPRDFVKMFKDVDAAVDAEVVQFMNSMAKNNITYKDLVKSCYHVMQ 600
  .                560               .                580               .                600

```

601 YSCNPFAQPA<sup>601</sup>CPIFTQLFYRSL<sup>620</sup>LTILQDISLPICMCYENDN<sup>640</sup>NPGLGQSPPEWLK<sup>660</sup>GHYQTLC

661 TNFRSLAIDKGVLTAKVAVHGEPTCDLPDLDAALQGRVYGRRLPVRMSKVLMLCPRNI 720  
 . 680 . 700 . 720

721 KIKNRVVF<sup>721</sup>TGENAALQNSFIKSTTR<sup>740</sup>RENYIINGP<sup>760</sup>YMKFLNTYHKTLFPD<sup>780</sup>TKLSSLYLWHN<sup>780</sup>

781 FSRRRSVPVPSGASAE EYSDLALFVDGGSRAHEESNVIDVVPGNLVTYAKQRLNNAILKA 840  
 . 800 . 820 . 840

841 CGQTQFYISLIQGLVPRTQSVPARDYPHVLGTRAVESAAAYAEATSSLTATTVVCAATDC 900  
 . 860 . 880 . 900

901 LSQVCKARPVVTLPTINKYTG VNGNNQIFQAGNLGYFMGRGVDRNLLQAPGAGLRKQAG 960  
 . 920 . 940 . 960

961 GSSMRKKFVFATPTLGLTVKRRTQAATTYEIENIRAGLEAIISQKQEEDCVFDVVCNLDV 1020  
 . 980 . 1000 . 1020

1021 AMGEACASLTRDDAEYLLGRFSVLADSVLETLATIASSGIEWTAEAAARDFLEGVWGPGGA 1080  
 . 1040 . 1060 . 1080

1 -----AVDTGSGGGGQPHDTAPRGARKKO 24

1081 A Q D N F I S V A E P V S T A S Q A S A G L L L G G G Q G S G G R R K R R L A T V L P G L E V \* 1129

% Identity = 0.0 (0/1129)      % Homology = 0.0 (0/1129)      % Total = 0.0 (0/1129)

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BA-LF2.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 1129 aa MQGAQTSEDNLG ... RLATVLPGLEV\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 18.** Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BA-LF2, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1 MQGAQTSEDNLG SQSQPGPCGYIYFYPLATYPLREVATLGTGYAGHRCLTVPLLCGITVE 60
   .                20                .                40                .                60
-----
61 PGFSINVKALHRRPDPNCGLLRATSYHRDIYVFHNAHMPPIFEGPGLEALCGETREVFG 120
   .                80                .                100               .                120
-----
121 YDAYSALPRESSKPGDFFPEGLDPSAYLGAVAITEAFKERLYSGNLVAIPSLKQEVAVGQ 180
   .                140               .                160               .                180
-----
181 SASVRVPLYDKEVFPEGVPLRQFYNSDLSRCMHEALYTGLAQALRVRRVGKLVELLEKQ 240
   .                200               .                220               .                240
-----
241 SLQDQAKVAKVAPLKEFPASTISHPD SGALMIVDSAACELAVSYAPAMLEASHETPASLN 300
   .                260               .                280               .                300
-----
301 YDSWPLFADCEGPEARVAALHRYNASLAPHVSTQIFATNSVLYVSGVSKSTGQGKESLFN 360
   .                320               .                340               .                360
-----
361 SFYMTHTGLGTLQEGTWDPCRRPCFSGWGGPDVTGTNGPGNYAVEHLVYAASFSPNLLARY 420
   .                380               .                400               .                420
-----
421 AYYLQFCQGQKSSLTPVPETGSYVAGAAASPMCSLCEGRAPAVCLNTLFFRLRDRFPPVM 480
   .                440               .                460               .                480
-----
481 STQRRDPYVISGASGSYNETDFLGNFLNFIDKEDDGQRPDDEPRYTYWQLNQNLLERLSR 540
   .                500               .                520               .                540
-----
541 LGIDAEGKLEKEPHGPRDFVKMFKDVDAAVDAEVVQFMNSMAKNNITYKDLVKSCYHVMQ 600
   .                560               .                580               .                600

```





9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BA-LF1.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ\*

Protein sequence 221 aa MNLAIALDSPHP ... LASLTDFLKSL\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 19.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BA-LF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .           40           .           60
1  MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEY 60
   M  +  +  L A  +  P           L  + P++ R A+ +  VF+ +  E  +
1  MNLAIALDSPHPGL=ASYTILPRPFYHISLKPVSWPDETMRPAKSTDSVFVRTPV==EAW 57
      .           20           .           40           .
      .           80           .           100          .
61  VQRTFGVPRRQRAIDKRQRASVAGAGAHHLGGSSATPVQQAQAAASAGTGAL-ASSAPS 119
   V  +  +  +  RA  A           A  +  +  A+           L A  A
58  VAPSPDDKVAESSYLMFRAMYAVFTRDEKDLPLPALVLCRLIKASLRKDRKLYAELACR 117
   60           .           80           .           100          .
   120          .           140          .           160          .
120  TAVAQSATPSVSSSISSLRAATSGATAAASAAAAMD--TGSGGGGQPHDTAPRGARKKQ* 177
   TA           V  IS LRA  +  S           V  T           D           A
118  TADIGGKDTHVRLIISVLRVYNDHYDYSRLRVVLCYTVVFAVRNYLDDHKSAAFVLGA 177
   120          .           140          .           160          .

-----

178  IAHYLALYRRLWFARLGGMPRLRRQFPVTWALASLTDFLKSL*                221
   180          .           200          .           220

```

% Identity = 14.7 (33/224)      % Homology = 8.0 (18/224)      % Total = 22.8 (51/224)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BA-LF1.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEEELLNKRVA\*

Protein sequence 221 aa MNLAIALDSPHP ... LASLTDFLKSL\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 20.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BA-LF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .           40           .           60
1  MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPFFSH 60
  M  A  +      G  A+      +  +  +  +P  +  T  +      P  +
1  MNLAIALDSPHP=GL=ASYTILPRPFYHISLKPVSWPDETMRPASTDSVFVRT=PVEAW 57
      .           20           .           40           .
61 QAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPPPPAYF 120
  A  +      +  +      Y      P      L  +      +  A
58 VAPSPDDDKVAESSYLMFRAMYAVFTRDEKDLPLPALVLCRLIKASLRKDRKLYAELACR 117
  60      .           80           .           100           .
121 GLPGLFGPPPPCLLTDSHLRADYVPAPSRSNKRKRDPEDDEEGGLFPGEDATLYRKDI 180
      G      L  S  LRA Y      +  R  R      +  D      +
118 TADIG=GKDTHVRLII=SVLRVYNDHYDYS=RLRVVLCYTVVFAVRNYLDDHKSAAFV 174
  120      .           140           .           160           .
181 AGLSKSVNELQHTLQALRRETLSTYGHGTGVGYCPQQGPCYTHSGPYGFGPHQSYEVPRYVP 240
      G      L  L  R  +      T
175 LGAIAHYLALYRRLWFARLGGMPRSLRRQFPVTWALASLTDFLKSL*===== 221
  180      .           200           .           220           .
241 HPPPPPTSHQAAQAQPPPPGTQAPEAHCAESTIPEAGAAGNSGPREDTNPQQPTTEGHH 300
      .           260           .           280           .           300
=====
      .           320           .           340
301 RGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEEELLNKRVA* 346
=====

```

% Identity = 7.8 (27/346)      % Homology = 5.2 (18/346)      % Total = 13.0 (45/346)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BA-LF1.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein sequence 221 aa MNLAIALDSPHP ... LASLTDFLKSL\*

Method: Diagonals (BLOSUM62)  
 Layout: Standard  
 Block Length ≤: 6-aa  
 Mismatch penalty: Smaller (1)  
 Gap penalty: Medium (2)  
 Display: BLOSUM62

Alignment 21. Comparison of the amino acid sequence of  
 SEQ ID NO:5 with the amino acid sequence, BA-LF1, encoded  
 by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .
      20
1 ---AVDTGSGGGGQPHDTA-PRGARKKQ----- 24
    A+  S  G  T  PR
1 MNLAIALDSPHPGLASYTILPRPFYHISLKPVSWPDETMRPAKSTDSVFVRTPVEAWVAP 60
      .
      20
      40
      60
-----
61 SPPDDKVAESSYLMFRAMYAVFTRDEKDLPLPALVLCRLIKASLRKDRKLYAELACRTAD 120
      .
      80
      100
      120
-----
121 IGGKDTHVRLIISVLRAVYNDHYDYSRLRVVLCYTVVFAVRNYLDDHKSAAFVLGAIAH 180
      .
      140
      160
      180
-----
181 YLALYRRLWFARLGGMPRSLRRQFPVTWALASLTDFLKSL* 221
      .
      200
      220
  
```

% Identity = 2.7 (6/221) % Homology = 0.5 (1/221) % Total = 3.2 (7/221)

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BA-LF1.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 221 aa MNLAIALDSPHP ... LASLTDFLKSL\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 22.** Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BA-LF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .
1 ST-AVAQ-SATPSVSSSISSLRAATSGATAAA----- 30
   A+A S P ++S R +
1 MNLAIALDSPHPGLASYTILPRPFYHISLKPVSWPDETMRPASTDSVFVRTPVEAWVAP 60
      .           20           .           40           .           60
-----
61 SPPDDKVAESSYLMFRAMYAVFTRDEKDLPLPALVLCRLIKASLRKDRKLYAELACRTAD 120
      .           80           .           100           .           120
-----
121 IGGKDTHVRLIISVLRAVYNDHYDYSRLRVVLCYTVVFAVRNYLDDHKSAAFVLGAIAH 180
      .           140           .           160           .           180
-----
181 YLALYRRLWFAFLGGMPRSLRRQFPVTWALASLTDFLKSL* 221
      .           200           .           220

```

% Identity = 2.7 (6/221) % Homology = 1.8 (4/221) % Total = 4.5 (10/221)

///

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BA-RF1.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ\*

Protein sequence 222 aa MARFIAQLLLLL ... HGVYVSGYLSQ\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 23.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, Ba-RF1, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .           40           .
1  MARRLPKPTLQGRLEADFPDSPL-LPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEE 59
   MAR + + L      A      L + L      + E + S+      +
1  MARFIAQLLLLLASCVAAGQAVTAFLGERVTLTSYWRRVSLGPEIEVSWFKLGPGEEQVLI 60
      .           20           .           40           .           60
60  YVQRTFGVPRRQRAIDKRQRASVAGAGAHHLGGSSATPVQQAQAAASAGTGALASSAPS 119
      +           A           + +           T           S
61  GRMHHDVIFIEWPFRGFFDIHRSANTFFLVVTAANISHDGNILCRMKLGETEVTKQEHLS 120
      .           80           .           100          .           120
120 TAVAQSATP-SVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ*- 177
      + + S S      T T A      V      G +P TA G      K+
121 VVKPLTSLVHSERSQFPDFSVLTVTCTVNAFPHPHVQWLMPEGVEPAPTAANGGVMKEKD 180
      .           140          .           160          .           180

```

```

-----
181 GSLSVAVDLSLPKPWHLPVTCVKGNDKEEAHGVYVSGYLSQ*                222
      .           200          .           220

```

% Identity = 10.8 (24/222)      % Homology = 6.3 (14/222)      % Total = 17.1 (38/222)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BA-RF1.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA\*

Protein sequence 222 aa MARFIAQLLLLA ... HGVIYVSGYLSQ\*

Method: Diagonals (BLOSUM62) Alignment 24. Comparison of the amino acid sequence  
 Layout: Standard encoded by the nucleotide sequence of SEQ ID NO:3  
 Block Length ≤: 6-aa (SEQ ID NO:4) with the amino acid sequence, Ba-RF1,  
 Mismatch penalty: Smaller (1) encoded by the nucleotide sequence of Fig. 2 of Bankier et al.  
 Gap penalty: Medium (2)  
 Display: BLOSUM62

```

      .           20           .           40           .           60
1  MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPPFSH 60
=====
      .           80           .           100          .           120
61 QAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAYF 120
      .                                     A F
1  =====MARF 4
      .           140          .           160           .           180
121 GLPGLFGPPPPCLLTDSHLRADYVPAPSRSNKRKRDPEDDEEGGLFPGEDATLYRKDI 180
      L           +       + V   S   +       PE +       L PGE+   L   +
.5 IAQLLLLASCVAAGQAVTAFLGERVTLTSYWRRVSLGPEIEVSWFKLGPGEEQVLIGRMH 64
      .           20           .           40           .           60
181 AGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYVP 240
      +       +       R       +       G   Y       G       E       V
65 HDVIFIEWPFRGFFDIHRSANTFFLVVTAANISHDG=NYLCRMKLGETEVTKQE=HLSV= 121
      .           80           .           100          .           120
241 HPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGHH 300
      P   S   +   ++Q   P   +       V       P       G   E   P       G
122 =VKPLTSLVHSERSQ=FPDFSVLTVTCTVNAFPHPHVQWLMPEG=VE=PAP=TAANGGVM 176
      .           140          .           160           .           180
301 RGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA* 346
      + K   + A   ++ K   P       +       +       *
177 KEKDGSLSVAVDLSLPKPWHLPVTCVGKNDKEEAHGVIYVSGYLSQ* 222
      180           .           200           .           220

```

% Identity = 9.5 (33/346)      % Homology = 5.5 (19/346)      % Total = 15.0 (52/346)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BA-RF1.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ  
Protein sequence 222 aa MARFIAQLLLLLA ... HGYYVSGYLSQ\*

Method: Diagonals (BLOSUM62)      **Alignment 25.** Comparison of the amino acid sequence of  
Layout: Standard      SEQ ID NO:5 with the amino acid sequence, Ba-RF1, encoded  
Block Length ≤: 6-aa      by the nucleotide sequence of Fig. 2 of Bankier et al.  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

```
-----  
1 MARFIAQLLLLLASCVAAGQAVTAFLGERVTLTSYWRRVSLGPEIEVSWFKLGPGEEQVLI 60  
      .                20      .                40      .                60  
-----  
61 GRMHHDVIFIEWPFRGFFDIHRSANTFFLVVTAANISHDGNILCRMKLGETEVTKEHLS 120  
      .                80      .                100      .                120  
-----  
121 VVKPLTSLVHSERSQFPDFSVLTVTCTVNAFPHPHVQWLMPEGVEPAPTAANGGVMKEKD 180  
      .                140      .                160      .                180  
      .                20  
1  -----AVDTGSGGGGQPHDTAPRGARKKQ 24  
      G      + H      G      +  
181 GSLSVAVDLSLPKPWHLPVTCVGKNDKEEAHGYYVSGYLSQ* 222  
      .                200      .                220  
-----  
% Identity = 1.4 (3/222)      % Homology = 0.9 (2/222)      % Total = 2.3 (5/222)
```

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BA-RF1.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 222 aa MARFIAQLLLLLA ... HGVIYVSGYLSQ\*

Method: Diagonals (BLOSUM62)  
 Layout: Standard  
 Block Length ≤: 6-aa  
 Mismatch penalty: Smaller (1)  
 Gap penalty: Medium (2)  
 Display: BLOSUM62

Alignment 26. Comparison of the amino acid sequence of  
 SEQ ID NO:6 with the amino acid sequence, Ba-RF1, encoded  
 by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .          20          .
1 STAV-AQSATPSVSSSISSLRAATSGATAAA----- 30
      AQ  +  +      A  G
1 MARFIAQLLLLLASCVAAGQAVTAFLGERVTLTSYWRRVSLGPEIEVSWFKLGPGEEQVLI 60
      .          20          .          40          .          60
-----
61 GRMHHDVIFIEWPFRGFFDIHRSANTFFLVVTAANISHDGNLYLCRMKLGETEVTKEHLS 120
      .          80          .          100          .          120
-----
121 VVKPLTSLVHSERSQFPDFSVLTVTCTVNAFPHPHVQWLMPEGVEPAPTAANGGVMKEKD 180
      .          140          .          160          .          180
-----
181 GSLSVAVDLSLPKPWHLPVTCVGKNDKEEAHGVIYVSGYLSQ* 222
      .          200          .          220
    
```

% Identity = 1.8 (4/222)      % Homology = 0.9 (2/222)      % Total = 2.7 (6/222)

///



9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BN-LF2a,b.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ\*

Protein sequence 163 aa MVHVLERALLEQ ... LSLRCELGWCG\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 27.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BN-LF2a,b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .           40           .           60
1  MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPNDVVFREAQRSYLVFLTSQFCYEEY 60
   M  L +  L+ +  A                               P DV R  R  LV L  F
1  MVHVLERALLEQQSSACGLPGSSTETRPSHPCPEDP=DVSRL==RLLLVLVCVLFGLLCL 57
      .           20           .           40           .
61 VQRTFGVPRRQRAIDKRQRASVAGAGAHHLGGSSATPVQQAQAAASAGTGALASSAPST 120
   +           R  R +           +           + A +           T  S
58 LLI*EATMRPGRPLAGFYATLRRSFRMSKRSKNKAKKERVVPVEDRPP=TPMPTSQRLIR 116
   60           .           80           .           100           .
121 AVAQSATPSVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ* 177
     A           +           R           S           D  S
117 RNALGGGVRPDAEDCIQRFHPLEPALGVSTKNF=DLLSLRCELGWCG*===== 163
     120           .           140           .           160

```

% Identity = 13.0 (23/177)      % Homology = 5.1 (9/177)      % Total = 18.1 (32/177)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BN-LF2a,b.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCELLNKRVA\*

Protein sequence 163 aa MVHVLERALLEQ ... LSLRCELGWCG\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 28.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BN-LF2a,b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .      .      20      .      40      .
1 MLSGNAGEGATACGGSAAAGQDLISVP-RNTFMTLLQTNLDNKPPRQTPLPYAAPLPFFS 59
  M+   E A   S+A G   S   R +   ++   L   L
1 MVHVL==ERALLEQQSSACGLPGSSTETRPSPHPCPEDPDVSRLRLLLVVLCVLFGLLCLL 58
      .      .      20      .      40      .
60 HQAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAY 119
    AT   P AG A   F   A +   T P
59 LI*EATMRPGRPLAGFYATLRRSFRMSKRSKNKAKKERVVEDRPP=TPMPTSQRLIRR 117
60      .      80      .      100      .
120 FGLPGLFGPPPPCLLTDSHLRADYVPAPSRSNKRKRDPEDDEEGGLFPGEDATLYRKD 179
    L G P   +   L   +   +   E   G
118 NALGGGVRPDAEDCIQRFHPLPALGVSTKNFDLLSLRCELGWCG*===== 163
120      .      140      .      160
180 IAGLSKSVNELQHTLQALRRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYV 239
=====

240      .      260      .      280      .
240 PHPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGH 299
=====

300      .      320      .      340
300 HRGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCELLNKRVA* 346
=====

```

% Identity = 7.5 (26/347)      % Homology = 2.6 (9/347)      % Total = 10.1 (35/347)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BN-LF2a,b.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein sequence 163 aa MVHVLERALLEQ ... LSLRCELGWCG\*

Method: Diagonals (BLOSUM62)  
 Layout: Standard  
 Block Length ≤: 6-aa  
 Mismatch penalty: Smaller (1)  
 Gap penalty: Medium (2)  
 Display: BLOSUM62

**Alignment 29.** Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BN-LF2a,b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20
1 AVDTGSGGGGQPHDTAPRGARKKQ----- 24
  V       +   +A
1 MVHVLERALLEQQSSACGLPGSSTETRP SHPCPEDPDVSRRLRLLLVVL CVLFGLLC LLLI 60
      .           20           .           40           .           60

-----

61 *EATMRPGRPLAGFYATLRRSFRMSKRSKNKAKKERVPVEDRPPTPMPTSQRLIRRNAL 120
      .           80           .           100           .           120

-----

121 GGGVRPDAEDCIQRFHPLEPALGVSTKNFDLLSLRCELGWCG* 163
      .           140           .           160

% Identity = 1.2 (2/163)    % Homology = 1.2 (2/163)    % Total = 2.5 (4/163)
    
```

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BN-LF2a,b.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 163 aa MVHVLERALLEQ ... LSLRCELGWCG\*

Method: Diagonals (BLOSUM62)  
 Layout: Standard  
 Block Length ≤: 6-aa  
 Mismatch penalty: Smaller (1)  
 Gap penalty: Medium (2)  
 Display: BLOSUM62

Alignment 30. Comparison of the amino acid sequence of  
 SEQ ID NO:6 with the amino acid sequence, BN-LF2a,b, encoded  
 by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .
1 STAVAQSATPSVSSSISSIRAAATSGATAAA----- 30
      V + A      SS  L  +++  +
1 MVHVLERALLEQQSSACGLPGSSTETRPSPHPCPEDPDVSRRLRLLLVVLGVLFGLLCLLLI 60
      .           20           .           40           .           60
-----
61 *EATMRPGRPLAGFYATLRRSFRMSKRSKNKAKKERVPVEDRPPTPMPTSQRLIRRNAL 120
      .           80           .           100           .           120
-----
121 GGGVRPDAEDCIQRFHPLEPALGVSTKNFDLLSLRCELGWCG* 163
      .           140           .           160

```

% Identity = 3.1 (5/163)      % Homology = 3.1 (5/163)      % Total = 6.1 (10/163)

///

9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BN-LF1b.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ\*

Protein sequence 269 aa VLGIWIYLLLEML ... PHGPVQLSYYD\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 31.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BN-LF1b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1 VLGIWIYLLLEMLWRLGATIWQLLAFFLAFFLDLILLIIALYLQQNWWTLLVDLLWLLFL 60
      .                20                .                40                .                60
1 -----MARRLPKPTLQGRLEADFPDSPLLPKFQ 28
      .                .                .                .                .                .
61 AILIWMYYHGORHSDEHHHDDSLPHPQQATDDSGHESDSNSNEGRHLLVSGAGDGPPLC 120
      .                .                .                .                .                .
      .                40                .                60                .                80                120
29 ELNQNNLPNDVFREARSYLVFLTSQFCYEEYVQRTFGVPRRQRAIDKRQRASVAGAGAH 88
      N                Q                +                P                Q                D
121 SQNLGAPGGGPDNGPQDPDNTDDNGPQDPDNTDDNGPHDPLPQ=DPDNTDDNGPQDPDNT 179
      .                .                .                .                .                .
      .                100                .                120                .                140
89 AHLGGSSATPVQQAQAAASAGTGALASSAPSTAVAQSATPSVSSSISSLRAATSGATAAA 148
      G                P                + +A + G G                +                P + +                + SG
180 DDNGPHDPLPHSPSDSAGNDG=GPPQLTEEEVENKGGDQGPPLMTDGGGGHSHDSGHGGGD 238
      180                .                200                .                220                .
      .                160                .
149 SAAAAVDTG-SGGGGQPHD-TAPRGARKKQ* 177
      + G SG GG D P                *
239 PHLPTLLLGSSSGSGGDDDDPHGPVQLSYYD* 269
      240                .                260

```

% Identity = 8.9 (24/271) % Homology = 4.4 (12/271) % Total = 13.3 (36/271)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BN-LF1b.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA\*

Protein sequence 269 aa VLGIWIYILLEML ... PHGPVQLSYYD\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 32.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BN-LF1b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .           40           .           60
1  MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPFFSH 60

=====

      .           80           .           100          .           120
61 QAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAYF 120
      V       Y       +   G       AF
1  =====VLGIWIY=LLEMLWRLGATIWQLLAFFLAFFLDLILLIIIALYLQ 44
      .           20           .           40
121 GLPGLFGPPPPCLLTDSHLRADYVPAPSRSNKRKRDPEDDEEGGGLFPGEDATLYRKDI 180
      L       LL   + L   Y       S++   D +
45 NWWTL=LVDLLWLLFLAILIWMYYHGQRHSDEHHHD=DSLPHPQQATDDSGHESDSNSN 102
      .           60           .           80           .           100
181 AGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYVP 240
      G   +       L   + L       G   PQ       +GP       +P
103 EGRHLLVSGAGDGPPLCSQNLGAPGGGPDNGPQDPDNTDDNGPQDPDNTDDNGPHDPLP 162
      .           120          .           140          .           160
241 -HPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGH 299
      P   +       G   P   H   ++S   + G   +   E+   Q
163 QDPDNTDDNGPQDPDNTDDNGPHDPLPHSPSDSAGNDGGPPQLTEEVENKGGDQGPPLMT 222
      .           180          .           200          .           220
300 HRGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA* 346
      G       S   G       PT       S   S       +       *
223 DGGGGHSHDSGHGGGDPHLPTLLLGSSSGGDDDDPHGPVQLSYYD* 269
      .           240          .           260

```

% Identity = 10.7 (37/347)      % Homology = 4.6 (16/347)      % Total = 15.3 (53/347)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BN-LF1b.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein sequence 269 aa VLGIWIYLLLEML ... PHGPVQLSYYD\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

Alignment 33. Comparison of the amino acid sequence of SEQ ID NO:5 with the amino acid sequence, BN-LF1b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1 VLGIWIYLLLEMLWRLGATIWQLLAFFLAFFLDLILLIIALYLLQNNWWTLLVDLLWLLLFL 60
      .                20                .                40                .                60
-----
61 AILIWMYYHGQRHSDEHHHDDSLPHPQQATDDSGHESDSNSNEGRHLLVSGAGDGPPLC 120
      .                80                .                100                .                120
-----
121 SQNLGAPGGGPDNGPQDPDNTDDNGPQDPDNTDDNGPHDPLPQDPDNTDDNGPQDPDNTD 180
      .                140                .                160                .                180
-----
181 DNGPHDPLPHSPSDSAGNDGGPPQLTEEEVENKGGDQGPPLMTDGGGGHSHDSGHGGGDPH 240
      .                200                .                220                .                240
      .                .                20
1 -----AVDTGSGGGGQP HDTAPRGARKKQ 24
      +GSGG                + +
241 LPTLLLGSSGSGGDDDDPHGPVQLSYYD* 269
      .                260

```

% Identity = 1.5 (4/269)      % Homology = 1.1 (3/269)      % Total = 2.6 (7/269)

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BN-LF1b.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 269 aa VLGIWIYILLEML ... PHGPVQLSYD\*

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 34.** Comparison of the amino acid sequence of SEQ ID NO:6 with the amino acid sequence, BN-LF1b, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

-----
1 VLGIWIYILLEMLWRLGATIWQLLAFFLAFFLDLILLIIALYLQQNWWTLLVDLLWLLLFL 60
   .                20                .                40                .                60
-----

61 AILIWMYYHGQRHSDEHHHDDSLPHPPQATDDSGHESDSNSNEGRHLLLVSGAGDGPPLC 120
   .                80                .                100                .                120
-----

121 SQNLGAPGGGPDNGPQDPDNTDDNGPQDPDNTDDNGPHDPLPQDPDNTDDNGPQDPDNTD 180
   .                140                .                160                .                180

1 -----ST 2

181 DNGPHDPLPHSPSDSAGNDGGPPQLTEEEVENKGGDQGPPLMTDGGGGHSHDSGHGGGDPH 240
   .                200                .                220                .                240
   .                20                .
3 AVAQSATPSVS-SSISSLRAATSGATAAA 30
   S S +
241 LPTLLLGSSSGSGGDDDDPHGPVQLSYD* 269
   .                260

% Identity = 0.7 (2/269)    % Homology = 0.4 (1/269)    % Total = 1.1 (3/269)

```

///



9310-13DVCTDV SEQ ID NO 2.xprt x Bankier et al. BN-LF1a.xprt => Protein Alignment

Protein sequence 177 aa MARRLPKPTLQG ... DTAPRGARKKQ\*

Protein sequence 144 aa MEHDLERGPPGP ... LGIVLFIFGCLL

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 35.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:1 (SEQ ID NO:2) with the amino acid sequence, BN-LF1a, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .           40           .           60
1  MARRLPKPTLQGRLEADFPDSPLLPKFQELNQNNLPNDVFREAQRSYLVFLTSQFCYEEY 60
   M  L  +      R      P      L      L      +      S      Y
1  MEHDLERGPPGPRRPPRGPPPLSSSLGLALLLLL=LALLFWLYIVMSDWTGGALLVLYSFA 59
      .           20           .           40           .
61 VQRTFGVPRRQRAIDKRQRASVAGAGAHHLGGSSATPVQQAQAAASAGTGALASSAPST 120
   +      +      I +R      GA      L      T      S+ + AL+ +P T
60 LMLIIIIILIIIF==IFRRDLLCPLGALCILLMSKYYTLCPTPPFPYSSFSNALSPSPVT 117
   60           .           80           .           100           .
      .           140           .           160           .
121 AVAQSATPSVSSSISSLRAATSGATAAASAAAAVDTGSGGGGQPHDTAPRGARKKQ* 177
   +      A  ++      L
118 LLLI=ALWNLHGQALFLGIVLFIFGCLL===== 144
   120           .           140

```

% Identity = 11.9 (21/177)      % Homology = 6.8 (12/177)      % Total = 18.6 (33/177)

///

9310-13DVCTDV SEQ ID NO 4.xprt x Bankier et al. BN-LF1a.xprt => Protein Alignment

Protein sequence 346 aa MLSGNAGEGATA ... FCEELLNKRVA\*

Protein sequence 144 aa MEHDLERGP PGP ... LGIVLFIFGCLL

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 36.** Comparison of the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:3 (SEQ ID NO:4) with the amino acid sequence, BN-LF1a, encoded by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .           40           .           60
1  MLSGNAGEGATACGGSAAAGQDLISVPRNTFMTLLQTNLDNKPPRQTPLPYAAPLPFFSH 60
  M      G      G      L      S      +      LL      L      +      A      L      +S
1  MEHDLER=GPPGPRRPPR=GPPLSSSLGLALLLLLLLALLFWLYIVMSDWTGGALLVLYSF 58
      .           20           .           40           .
61  QAIATAPSYGPGAGAVAPAGGYFTSPGGYYAGPAGGDPGAFLAMDAHTYHPHPHPPPAYF 120
  +
59  ALMLIIIIIIIFIFRRDLLCPLGALCILLMSKYYTLCPTPPFPYSSFSNALSPVTL 118
60      .           80           .           100           .
121 GLPGLFGPPPPCLLT TDSHLRADYVPAPSRSNKRKRDP EEEGGGLFPGEDATLYRKDI 180
  L      L+      L      +
119 LLIALWNLHGQALFLGIVLFIFGCLL===== 144
120      .           140           .           160           .           180
181 AGLSKSVNELQHTLQALRRETLSYGHTGVGYCPQQGPCYTHSGPYGFQPHQSYEVPRYVP 240
=====
      .           260           .           280           .           300
241 HPPPPPTSHQAAQAQPPPPGTQAPEAHCVAESTIPEAGAAGNSGPREDTNPQQPTTEGHH 300
=====
      .           320           .           340
301 RGKKLVQASASGVAQSKEPTTPKAKSVSAHLKSIFCEELLNKRVA* 346
=====

```

% Identity = 4.3 (15/346)      % Homology = 2.3 (8/346)      % Total = 6.6 (23/346)

///

9310-13DVCTDV SEQ ID NO 5.xprt x Bankier et al. BN-LF1a.xprt => Protein Alignment

Protein sequence 24 aa AVDTGSGGGGQP ... HDTAPRGARKKQ

Protein sequence 144 aa MEHDLERGPPGP ... LGIVLFIFGCLL

Method: Diagonals (BLOSUM62)  
Layout: Standard  
Block Length ≤: 6-aa  
Mismatch penalty: Smaller (1)  
Gap penalty: Medium (2)  
Display: BLOSUM62

**Alignment 37.** Comparison of the amino acid sequence of  
SEQ ID NO:5 with the amino acid sequence, BN-LF1a, encoded  
by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20
1 AVDTGSGGGGQPHDTAPRGARKKQ----- 24
      G   P   PRG
1 MEHDLERGPPGP=PRGPPLSSSLGLALLLLLALLFWLYIVMSDWTGGALLVLYSFA 59
      .           20           .           40           .

-----

60 LMLIIIIIIIFIFRRDLLCPLGALCILLMSKYYTLCPTPPFPYSSFSNALSPVLTLL 119
60           .           80           .           100           .

-----

120 LIALWNLHGQALFLGIVLFIFGCLL 144
120           .           140

% Identity = 3.4 (5/145)    % Homology = 0.0 (0/145)    % Total = 3.4 (5/145)
```

///

9310-13DVCTDV SEQ ID NO 6.xprt x Bankier et al. BN-LF1a.xprt => Protein Alignment

Protein sequence 30 aa STAVAQSATPSV ... LRAATSGATAAA

Protein sequence 144 aa MEHDLERGPPGP ... LGIVLFIFGCLL

Method: Diagonals (BLOSUM62)  
 Layout: Standard  
 Block Length ≤: 6-aa  
 Mismatch penalty: Smaller (1)  
 Gap penalty: Medium (2)  
 Display: BLOSUM62

**Alignment 38.** Comparison of the amino acid sequence of  
 SEQ ID NO:6 with the amino acid sequence, BN-LF1a, encoded  
 by the nucleotide sequence of Fig. 2 of Bankier et al.

```

      .           20           .
1 STAVAQSATPSVSSSISSLRAATSGATAAA----- 30
      +   P           ++S   A
1 MEHDLERGPPGPRRPPRGPPPLSSSLGLALLLLLLLALLFWLYIVMSDWTGGALLVLYSFAL 60
      .           20           .           40           .           60

-----

61 MLIIIIILIIIFIFRRDLLCPLGALCILLMSKYYTLCPTPPFPYSSFSNALSPVLTLL 120
      .           80           .           100           .           120

-----

121 IALWNLHGQALFLGIVLFIFGCLL 144
      .           140
  
```

% Identity = 2.1 (3/144)      % Homology = 2.1 (3/144)      % Total = 4.2 (6/144)

///